

ORIGIN

Query Match 94.6%; Score 969; DB 6; Length 4487;  
Best Local Similarity 99.0%; Pred. No. 8.9e-227; Mismatches 5; Indels 5; Gaps 4;  
Matches 1017; Conservative 0;

Qy 1 ACCGCTCGTCCGCGCGCTTGAGGCGCGCGGCGCGCGCAATTCGTGCGCGCGCGG 60  
Db 4 ACCGCTCGTCCG-CGCGCTTGAGGCGCGCGGAG-GCGCGCAATTCGTGCGCGCGG 61

Qy 61 GGGGCGGCC--TCCGCGCATCTTCGCGCGCAACAAGACTACAGAGAGGAGGCGGTG 118  
Db 62 GGGGCGGCCATCCGCGCAGTCTTCGCGCGCAACAAGACTACAGAGAGGAGGCGGTG 121

Qy 119 GGATGCGCGTCCGCGCGCGCGAGTACAAAGCGGCGACCTGGTCTTCGCAAGATGA 178  
Db 122 GGATGCGCGTCCGCGCGCGCGAGTACAAAGCGGCGACCTGGTCTTCGCAAGATGA 181

Qy 179 AGGGCTACCGCACTGGCGCGCGCGGATGATGAATCCAGAGGCGGTGTGAAGCCTC 238  
Db 182 AGGGCTACCGCACTGGCGCGCGCGGATGATGAATCCAGAGGCGGTGTGAAGCCTC 241

Qy 239 CAGCAAAACAAGTATCTATCTCTTTTGGCACCCATGAATCTGATTTCTAGTCCCA 298  
Db 242 CAGCAAAACAAGTATCTATCTCTTTTGGCACCCATGAATCTGATTTCTAGTCCCA 301

Qy 299 AAGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACGAAAG 358  
Db 302 AAGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACGAAAG 361

Qy 359 GATTTAACAGAGGATTTGGGAAATAGAAATAAACCCAGGAGTAAAGTTTACTGGCTAC 418  
Db 362 GATTTAACAGAGGATTTGGGAAATAGAAATAAACCCAGGAGTAAAGTTTACTGGCTAC 421

Qy 419 AGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGGAAGAGTGAATACTGCGAGTG 478  
Db 422 AGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGGAAGAGTGAATACTGCGAGTG 481

Qy 479 CAAGCAGTGGGAAGAGTGTAGAGTAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 538  
Db 482 CAAGCAGTGGGAAGAGTGTAGAGTAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 541

Qy 539 AAAAGCAGGCTCAAAACGGAAGTCAATATCTTCAAGAAATCCTCTAAACAGTCCC 598  
Db 542 AAAAGCAGGCTCAAAACGGAAGTCAATATCTTCAAGAAATCCTCTAAACAGTCCC 601

Qy 599 GGAATCTCAGAGATGAAGATGACAAAGCTGCAAGAGAGGAAACAAAGAGCT 658  
Db 602 GGAATCTCAGAGATGAAGATGACAAAGCTGCAAGAGAGGAAACAAAGAGCT 661

Qy 659 CTGAGGCTGAGATCGGCGCAACGACACAGAAACCAACTTCAGACTTCGAGAAACCA 718  
Db 662 CTGAGGCTGAGATCGGCGCAACGACACAGAAACCAACTTCAGACTTCGAGAAACCA 721

Qy 719 GTGAAGGAGCTTAACCTACCAATATGAATGCTGCATATTAAGAGAAACCAAGAGTT 777  
Db 722 GTGAAGGAGCTTAACCTACCAATATGAATGCTGCATATTAAGAGAAACCAAGAGTT 781

Qy 778 ATATGTTGGTGTCTTAATATCTTGGATTTGATGAACCAACACATAGTCTCTTTGT 837  
Db 782 ATATGTTGGTGTCTTAATATCTTGGATTTGATGAACCAACACATAGTCTCTTTGT 841

Qy 838 CATTTGACAGAACCCAGTTGTATGTACATATTCATATTCCTCTGTTGTTTCGG 897  
Db 842 CATTTGACAGAACCCAGTTGTATGTACATATTCATATTCCTCTGTTGTTTCGG 901

Qy 898 GGGAAAGACATTTTAGCCTTTTTTAAAGTTTACTGATTTAATTTTCAATTTTGGTTG 957  
Db 902 GGGAAAGACATTTTAGCCTTTTTTAAAGTTTACTGATTTAATTTTCAATTTTGGTTG 961

Qy 958 CATGAAGTTGCCCTTAACCACTAAGGATTAACAGATTTTGGCAGACTTATACATGTC 1017  
Db 962 CATGAAGTTGCCCTTAACCACTAAGGATTAACAGATTTTGGCAGACTTATACATGTC 1021

Qy 1018 TAGGATC 1024  
Db 1022 TAGGATC 1028

RESULT 9  
AR240312  
LOCUS AR240312 1664 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 68 from patent US 6468758.  
ACCESSION AR240312  
VERSION AR240312.1 GI:27285392  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1664)  
AUTHORS Benson,D.R., Lodes,M.J., Mitcham,J.L. and King,G.B.  
TITLE Compositions and methods for ovarian cancer therapy and diagnosis  
JOURNAL Patent: US 6468758-A 68 22-OCT-2002;  
FEATURES Location/Qualifiers  
source 1..1664  
BASE COUNT 566 a 280 c 319 g 499 t  
ORIGIN /organism="unknown"

Query Match 80.9%; Score 828.4; DB 6; Length 1664;  
Best Local Similarity 98.7%; Pred. No. 2.4e-192;  
Matches 835; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 179 AGGGCTACCCGCACTGGCGCGCGCGGATTTGATGAATCCAGAGGCGCTGTGAAGCCTC 238  
Db 17 AGGGCGCGCCAGTGGCGGTAGCCCGGATTTGATGAATCCAGAGGCGCTGTGAAGCCTC 76

Qy 239 CAGCAAAACAAGTATCTATCTCTTTTGGCAACCATGAATCTGCAATTTCTAGTCCCA 298  
Db 77 CAGCAAAACAAGTATCTATCTCTTTTGGCAACCATGAATCTGCAATTTCTAGTCCCA 136

Qy 299 AAGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACGAAAG 358  
Db 137 AAGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACGAAAG 196

Qy 359 GATTTAACAGAGGATTTGGGAAATAGAAATAAACCCAGGAGTAAAGTTTACTGGCTACC 418  
Db 197 GATTTAACAGAGGATTTGGGAAATAGAAATAAACCCAGGAGTAAAGTTTACTGGCTACC 256

Qy 419 AGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGGAAGAGTGAATACTGCGAGTG 478  
Db 257 AGGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGGAAGAGTGAATACTGCGAGTG 316

Qy 479 CAAGCAGTGGAGAGAGTGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAG 538  
Db 317 CAAGCAGTGGAGAGAGTGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAGAGTAG 376

Qy 539 AAAAGCAGGCTCAAAACGGAAGTCAATATCTTCAAGAAATCCTCTAAACAGTCCC 598  
Db 377 AAAAGCAGGCTCAAAACGGAAGTCAATATCTTCAAGAAATCCTCTAAACAGTCCC 436

Qy 599 GGAATCTCAGAGATGAAGATGCAAGAGTCAAGAGAGAGGAAACAAAGAGCT 658  
Db 437 GGAATCTCAGAGATGAAGATGCAAGAGTCAAGAGAGAGGAAACAAAGAGCT 496

Qy 659 CTGAGGCTGAGATCGGCGCAACGACACAGAAACCAACTTCAGACTTCGAGAAACCA 718  
Db 497 CTGAGGCTGAGATCGGCGCAACGACACAGAAACCAACTTCAGACTTCGAGAAACCA 556

Qy 719 GTGAAGGAGCTTAACCTACCAATATGAATGCTGCATATTAAGAGAAACCAAGAGTTA 778  
Db 557 GTGAAGGAGCTTAACCTACCAATATGAATGCTGCATATTAAGAGAAACCAAGAGTTA 616

Qy 779 TATGTTGGTGTCTTAATATCTTGGATTTGATGAACCAACACATAGTCTCTTTGTC 838  
Db 617 TATGTTGGTGTCTTAATATCTTGGATTTGATGAACCAACACATAGTCTCTTTGTC 676

priority-9/23/02



Db 482 CAAGCAGTGTAGGAAGAAGGTGATAGAGTAGAAGAGATGGAAGAAGGCAAAAGGAAGAATG 541  
Qy 539 AAAAAAGCAGGCTCAAAAACGGAAGAAAGTCATATACCTCAAAAGAAATCCTCTAAACAGTCCC 598  
Db 542 AAAAAAGCAGGCTCAAAAACGGAAGAAAGTCATATACCTCAAAAGAAATCCTCTAAACAGTCCC 601  
Qy 599 GGAATCTCAGGAGATGAAGATGACAAAGACTGCAAAAGAGAGAAACAAAAGACAGCT 658  
Db 602 GGAATCTCAGGAGATGAAGATGACAAAGACTGCAAAAGAGAGAAACAAAAGACAGCT 661  
Qy 659 CTGAGGCTGAGATGCGGCAACGACACAAAGAACACAACTTCAGACTTCGAGAACCA 718  
Db 662 CTGAGGCTGAGATGCGGCAACGACACAAAGAACACAACTTCAGACTTCGAGAACCA 721  
Qy 719 GTGAAGGAGCTAACTACCAATGAATGCTGCATATTAAGAGAAACCAAA-GAAGGTT 777  
Db 722 GTGAAGGAGCTAACTACCAATGAATGCTGCATATTAAGAGAAACCAAGAGGTT 781  
Qy 778 ATATGTTGTTGTTCTAATATCTTGGATTGATGAACCAACACATAGTCTCTTGTGT 837  
Db 782 ATATGTTGTTGTTCTAATATCTTGGATTGATGAACCAACACATAGTCTCTTGTGT 841  
Qy 838 CATTCAGAACCCAGTTTGTATGATACATATTCATATTCCTCTCTGTTGTTTGGG 897  
Db 842 CATTCAGAACCCAGTTTGTATGATACATATTCATATTCCTCTCTGTTGTTTGGG 901  
Qy 898 GGGAAAGACATTTTACGCTTTTAAAGTTACTGATTAATTTCAATGTTTGGTTG 957  
Db 902 GGGAAAGACATTTTACGCTTTTAAAGTTACTGATTAATTTCAATGTTTGGTTG 961  
Qy 958 CATGAAGTTGCCCTTAACCACTAAGGATTATCAAGATTTTGGCGAGCTTATACATGTC 1017  
Db 962 CATGAAGTTGCCCTTAACCACTAAGGATTATCAAGATTTTGGCGAGCTTATACATGTC 1021  
Qy 1018 TAGGATC 1024  
Db 1022 TAGGATC 1028

RESULT 7  
ABT06906  
ID ABT06906 standard; cDNA; 1664 BP.  
XX  
AC ABT06906;  
XX  
XX  
DT 07-NOV-2002 (first entry)  
XX  
DE Human ovarian cancer associated coding sequence SEQ ID NO: 68.  
XX  
KW Human; ovarian cancer; cancer; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX  
PN US2002076715-A1.  
XX  
PD 20-JUN-2002.  
XX  
XX  
PF 06-JUN-2001; 2001US-0876889.  
XX  
PR 23-SEP-1998; 98US-0159320.  
XX  
PR 08-FEB-1999; 99US-0246429.  
XX  
PR 16-SEP-1999; 99US-0397787.  
XX  
XX  
PA (BENS/) BENSON D R.  
PA (LODE/) LODES M J.  
PA (MITC/) MITCHAM J L.  
PA (KING/) KING G E.  
XX  
XX Benson DR, Lodes MJ, Mitcham JL, King GE;  
PI WPI; 2002-598720/64.  
XX  
XX

Composition for detecting and treating ovarian cancer, comprises a specific polypeptide, polynucleotide, T cell population, or antigen presenting cell -  
XX  
XX Example 1; Fig 16; 188pp; English.  
XX  
CC The present invention relates to a method of detecting the presence of ovarian cancer in a patient, involving detecting ovarian cancer associated polynucleotides. The method is not only used to detect the presence of cancer, preferably ovarian cancer in a patient, but also is used to stimulate and/or expand T cells specific for an ovarian tumour protein. The sequences can be used in vaccines used to treat cancer. The present sequence is an ovarian cancer associated coding sequence.  
XX  
SQ Sequence 1664 BP; 566 A; 280 C; 319 G; 499 T; 0 other;  
Query Match 80.9%; Score 828.4; DB 24; Length 1664;  
Best Local Similarity 98.7%; Pred. No. 3.6e-211;  
Matches 835; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
Qy 179 AGGGCTACCGCAGCTGGCGGCCCGGATTTGATGAACCTCCAGAGGGCGCTGTGAAGCCTC 238  
Db 17 AGGGCGGCGAGTGGCGGTAGCCCGATTTGATGAACCTCCAGAGGGCGCTGTGAAGCCTC 76  
Qy 239 CAGCAAAACAAGTATCTATCTTTTGGCACCCATGAAACTGCAATTTCTAGGTCCCA 298  
Db 77 CAGCAAAACAAGTATCTATCTTTTGGCACCCATGAAACTGCAATTTCTAGGTCCCA 136  
Qy 299 AGACCTTTTTCATATAAGGAGTACAAAGACAGTTTGGAAAGTCAAAACGGAAG 358  
Db 137 AGACCTTTTTCATATAAGGAGTACAAAGACAGTTTGGAAAGTCAAAACGGAAG 196  
Qy 359 GATTTAAACGAGATTTGGGAAATAGAAATAACCCAGAGTAAAGTTTACTGGCTACC 418  
Db 197 GATTTAAACGAGATTTGGGAAATAGAAATAACCCAGAGTAAAGTTTACTGGCTACC 256  
Qy 419 AGGCAATTCAGCAACAGAGCTCTTTCAGAAACTGAGGAGAGGTGGAATATCTGCAGATG 478  
Db 257 AGGCAATTCAGCAACAGAGCTCTTTCAGAAACTGAGGAGAGGTGGAATATCTGCAGATG 316  
Qy 479 CAAGCAGTGAAGAGAGGTGATAGAGTAGAAGAGATGGAAGGCAAAAGAAAGAAATG 538  
Db 317 CAAGCAGTGAAGAGAGGTGATAGAGTAGAAGAGATGGAAGGCAAAAGAAAGAAATG 376  
Qy 539 AAAAAAGCAGGCTCAAAACGGAAGAAAGTCAATATCTTCAAGAAATCCTCTAAACAGTCCC 598  
Db 377 AAAAAAGCAGGCTCAAAACGGAAGAAAGTCAATATCTTCAAGAAATCCTCTAAACAGTCCC 436  
Qy 599 GGAATCTCCAGGAGATGAAGATGACAAAGACTGCAAAAGAGAGGAAAAACAAAGCAGCT 658  
Db 437 GGAATCTCCAGGAGATGAAGATGACAAAGACTGCAAAAGAGAGGAAAAACAAAGCAGCT 496  
Qy 659 CTGAGGCTGAGATGCGGCAACGACACAAAGAAACAACTTCAGACTTCGAGAAACCA 718  
Db 497 CTGAGGCTGAGATGCGGCAACGACACAAAGAAACAACTTCAGACTTCGAGAAACCA 556  
Qy 719 GTGAAGGAGCTTAATACCATATGATGCTGCATATTAAGAGAAACCAAGAGGTTA 778  
Db 557 GTGAAGGAGCTTAATACCATATGATGCTGCATATTAAGAGAAACCAAGAGGTTA 616  
Qy 779 TATGTTTGGTTGTTCTAATATTTCTTGGATTGATATGAACCAACACATAGTCTCTTGTGTC 838  
Db 617 TATGTTTGGTTGTTCTAATATTTCTTGGATTGATATGAACCAACACATAGTCTCTTGTGTC 676  
Qy 839 ATTGACAGAAACCCAGTTTGTATGATACATATTCATATTCCTCTCTGTTGTTTGGGG 898  
Db 677 ATTGACAGAAACCCAGTTTGTATGATACATATTCATATTCCTCTCTGTTGTTTGGGG 736  
Qy 899 GGAAGAGACATTTTAGCCTTTTAAAGTTACTGATTAATTTCAATGTTTACTGTTGTC 958  
Db 737 GGAAGAGACATTTTAGCCTTTTAAAGTTACTGATTAATTTCAATGTTTACTGTTGTC 796  
Qy 959 ATGAAGTTGCCCTTAACCACTAAGGATTATCAAGATTTTGGCGAGACTTATACATGTC 1018

Db 797 ATGAAGTGGCTTACCACTAAGGATTATCAAGATTTTGGCGAGACTTATACATGTCT 856  
QY 1019 AGGATC 1024  
Db 857 AGGATC 862

RESULT 8  
ABT07105  
ID ABT07105 standard; cDNA; 1664 BP.  
XX  
AC  
XX  
XX  
DT 07-NOV-2002 (first entry)  
XX  
DE Human ovarian cancer associated coding sequence SEQ ID NO: 267.  
XX  
KW Human; ovarian cancer; cancer; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX  
PN US2002076715-A1.  
XX  
PD 20-JUN-2002.  
XX  
XX 06-JUN-2001; 2001US-0876889.  
XX  
PF 23-SEP-1998; 98US-0159320.  
PR 08-FEB-1999; 99US-0246429.  
PR 16-SEP-1999; 99US-0397787.  
XX  
XX (BENS/) BENSON D R.  
PA (LODE/) LODES M J.  
PA (MITC/) MITCHAM J L.  
PA (KING/) KING G E.  
XX  
XX Benson DR, Lodes MJ, Mitcham JL, King GE;  
XX WPI; 2002-598720/64.  
XX  
XX Composition for detecting and treating ovarian cancer, comprises a  
PT specific polypeptide, polynucleotide, T cell population, or antigen  
PT presenting cell -  
XX  
XX Example 1; Page 100-101; 189pp; English.  
XX  
XX The present invention relates to a method of detecting the presence of  
CC ovarian cancer in a patient, involving detecting ovarian cancer  
CC associated polynucleotides. The method is not only used to detect the  
CC presence of cancer, preferably ovarian cancer in a patient, but also is  
CC used to stimulate and/or expand T cells specific for an ovarian tumour  
CC protein. The sequences can be used in vaccines used to treat cancer. The  
CC present sequence is an ovarian cancer associated coding sequence.  
XX  
XX  
SQ Sequence 1664 BP; 566 A; 280 C; 319 G; 499 T; 0 other;

Query Match 80.9%; Score 828.4; DB 24; Length 1664;  
Best Local Similarity 98.7%; Pred. No. 3.6e-211;  
Matches 835; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 179 AGGGCTACCGCACTGGCGCGCGGATGATGAACCTCCAGAGGCGCTGTGAAGCCTC 238  
Db 17 AGGGCGCGCGAGTGGCGGTAGCCGGATGATGAACCTCCAGAGGCGCTGTGAAGCCTC 76

QY 239 CAGCAAAACAAAGTATCTCTCTTTTGGCACCCTCAAACTGCATTTCTAGGTCCCA 298  
Db 77 CAGCAAAACAAAGTATCTCTCTTTTGGCACCCTCAAACTGCATTTCTAGGTCCCA 136

QY 299 AAGACCTTTTCCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACCAACGGAAG 358  
Db 137 AAGACCTTTTCCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACCAACGGAAG 196

QY 359 GATTTAACGAAGGATTTGGGAAATAGAAAATAAACCAGGAGTAAAGTTTACTGGCTACC 418  
Db 197 GATTTAACGAAGGATTTGGGAAATAGAAAATAAACCAGGAGTAAAGTTTACTGGCTACC 256

QY 419 AGGCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAGGTGGAATATCTGCAGATG 478  
Db 257 AGGCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAGGTGGAATATCTGCAGATG 316

QY 479 CAAGCAGTGAGGAAGAAGTGTATAGTAGAAGAGATGGAAGGCGGCAAGAAAGAAATG 538  
Db 317 CAAGCAGTGAGGAAGAAGTGTATAGTAGAAGAGATGGAAGGCGGCAAGAAAGAAATG 376

QY 539 AAAAGCAGGCTCAAAACGGAAGAGTCAATATCTTCAAAGAAATCTCTTAAACAGTCCC 598  
Db 377 AAAAGCAGGCTCAAAACGGAAGAGTCAATATCTTCAAAGAAATCTCTTAAACAGTCCC 436

QY 599 GGAATCTCCAGGAGATGAAGATGCAAAAGACTGCAAAAGAGAGGAAACCAAAACAGCT 658  
Db 437 GGAATCTCCAGGAGATGAAGATGCAAAAGACTGCAAAAGAGAGGAAACCAAAACAGCT 496

QY 659 CTGAGGTTGGAGATGCGGGCAACGACACAAGAAACACAACCTTCAGACTTGCAGAAAACCA 718  
Db 497 CTGAGGTTGGAGATGCGGGCAACGACACAAGAAACACAACCTTCAGACTTGCAGAAAACCA 556

QY 719 GTGAAGGACCTTAACCTACATAATGAATGCTGCATATTAAGAGAAACCAACAAGAGTTA 778  
Db 557 GTGAAGGACCTTAACCTACATAATGAATGCTGCATATTAAGAGAAACCAACAAGAGTTA 616

QY 779 TATGTTGGTGTCTAATATTTCTTGGAATTTGATATGAACCAACATAGTCTCTGTTGTC 838  
Db 617 TATGTTGGTGTCTAATATTTCTTGGAATTTGATATGAACCAACATAGTCTCTGTTGTC 676

QY 839 ATTGACAGAACCCAGTTTGTATGTACATTTTCAATATTTCTCTCTGTTGTTGCGGG 898  
Db 677 ATTGACAGAACCCAGTTTGTATGTACATTTTCAATATTTCTCTCTGTTGTTGCGGG 736

QY 899 GGAAGAAGACATTTTAGCCTTTTAAAAAGTTTACTGATTTAATTTTCAATGTTGTTGTC 958  
Db 737 GGAAGAAGACATTTTAGCCTTTTAAAAAGTTTACTGATTTAATTTTCAATGTTGTTGTC 796

QY 959 ATGAAGTTGCCCTTAACCACTAAGGATTTCAAGATTTTGGCGAGACTTATACATGCT 1018  
Db 797 ATGAAGTTGCCCTTAACCACTAAGGATTTCAAGATTTTGGCGAGACTTATACATGCT 856

QY 1019 AGGATC 1024  
Db 857 AGGATC 862

RESULT 9  
ABX72784  
ID ABX72784 standard; cDNA; 1664 BP.  
XX  
AC  
XX  
XX 14-MAR-2003 (first entry)  
XX  
XX Human ovarian carcinoma antigen partial cDNA sequence #50.  
DE Human; cancer detection; ovarian carcinoma antigen; ovarian cancer;  
KW tumour antigen; tumour; OV2; OV3; OV6; OV9; OV10; OV12; OV14; OV17;  
KW OV18; OV23; OV24; OV27; OV41; OV54; OV57; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX US6468758-B1.  
XX  
XX 22-OCT-2002.  
XX  
XX 16-SEP-1999; 99US-0397787.  
XX  
XX 23-SEP-1998; 98US-0159320.  
PR 08-FEB-1999; 99US-0246429.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2004, 06:11:45 ; Search time 3909 Seconds  
(without alignments)  
10716.669 Million cell updates/sec

Title: US-09-787-328B-3  
Perfect score: 1024  
Sequence: 1 accgtctgcgcgcggctt.....acttatcatgtctaggatc 1024

Scoring table: IDENTITY NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues  
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sta.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_ot.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sta.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1013	98.9	1423	9	AF110642 Homo sapi
2	1013	98.9	1830	9	AF151900 Homo sapi
3	1012	98.8	1973	6	BD156109 Primer fo
4	1012	98.8	1973	9	AK001280 Homo sapi
5	1001	97.8	1427	9	AB029156 Homo sapi
6	990	96.7	1817	9	BC015483 Homo sapi
7	980	95.7	3805	6	BD131023 Human gto
8	969	94.6	4487	6	AR220855 Sequence
9	828.4	80.9	1664	6	AR240312 Sequence
10	828.4	80.9	1664	6	AR240511 Sequence
11	772.8	75.5	2750	10	AF389347 Rattus no
12	623.6	60.9	737	6	BD147622 Primer fo
13	547.8	53.5	612	10	AB029493 Mus muscu
14	425	41.5	1247	9	IR0783648 Homo sapi
15	363	35.4	13005	9	AL590240 Human DNA
16	299	29.2	2837	9	HS0801382 Homo sapi
17	299	29.2	148204	2	AC069400 Homo sapi
18	299	29.2	157624	2	AC087666 Homo sapi
19	299	29.2	182741	2	AC018910 Homo sapi
20	299	29.2	184113	9	AC024270 Homo sapi
21	263.6	25.7	334	6	BD054901 Sequence
22	195	19.0	2096	9	BC018991 Homo sapi
23	193.4	18.9	2376	6	AR081931 Sequence
24	193.4	18.9	2376	6	AX302495 Sequence
25	193.4	18.9	2376	6	E08546 DNA encodin
26	193.4	18.9	2376	9	HUMHDPF Human mRNA
27	192.4	18.8	113199	9	AC103876 Homo sapi
28	189	18.5	723	6	E08545 DNA encodin
29	188.6	18.4	869	6	AR081930 Sequence
30	187	18.3	1932	4	BT237996 Bos tauru
31	186	18.2	1752	10	AF251787 Mus muscu
32	185	18.1	1744	10	BC005713 Mus muscu
33	185	18.1	1763	10	BC021654 Mus muscu
34	184.2	18.0	711	10	AF448810 Rattus no
35	184.2	18.0	1571	10	AF389348 Rattus no
36	183.8	17.9	714	6	E13960 Mouse mRNA
37	183.8	17.9	1563	6	E13961 Mouse mRNA
38	183.8	17.9	1563	10	D83707 Mus musculu
39	182.6	17.8	1966	10	AF339083 Mus muscu
40	182.6	17.8	2559	10	AF339084 Rattus no
41	182.6	17.8	2831	10	AF339082 Mus muscu
42	181	17.7	965	10	BC002104 Mus muscu
43	181	17.7	1255	10	MMU308966 Mus muscu
44	181	17.7	1853	10	BC052177 Mus muscu
45	181	17.7	2928	10	MMU308965 Mus muscu

ALIGNMENTS

RESULT 1  
AF110642  
LOCUS AF110642 1423 bp mRNA linear PRI 11-MAY-2002  
DEFINITION Homo sapiens hepatoma-derived growth factor 2 (HDGF2) mRNA,  
complete cds.  
ACCESSION AF110642  
VERSION AF110642.1 GI:20530136  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1423)  
AUTHORS Yu,L., Fu,Q. and Tu,Q.  
TITLE Cloning of a novel human cDNA which is a homolog to mouse

hepatoma-derived growth factor (mHDGF) and termed hHDGF2  
Unpublished  
2 (bases 1 to 1423)  
Yu L., Zhang H.L., Fu, Q. and Zhao, Y.  
Direct Submission  
Submitted (04-DEC-1998) Lab of Human Gene Research, Institute of  
Genetics, 220 Handan Road, Shanghai 200433, China  
Location/Qualifiers  
1. .1423  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1. .1423  
/gene="HDGF2"  
125. .736  
/gene="HDGF2"  
/note="similar to Mus musculus hepatoma-derived growth  
factor"  
/codon\_start=1  
/product="hepatoma-derived growth factor 2"  
/protein\_id="AA027001.1"  
/db\_xref="GI:20530137"  
/translation="MARPRREYKAGDLVFAVMKGYPHWPARIDELPEGAVKPPANKY  
PIFFGTHETAFGLPKDLFPYKEYDKFGKSNKRKFNGLWEIENNPVGFYQAI  
QQSSSTEGEGNTADASSEEDGVEEDGKGRKNEKAGSKKSYTSKSSKQSR  
KSPGDEDDKDCKEENKSSSEGGDAGNDTRNTTSDLQKTSST"  
polyA\_signal  
1393. .1398  
/gene="HDGF2"  
BASE COUNT 464 a 282 c 320 g 357 t  
ORIGIN  
Query Match 98.9%; Score 1013; DB 9; Length 1423;  
Best Local Similarity 99.9%; Pred. No. 1.1e-237;  
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
Qy 1 ACCGCTGTCGCGCGGCTTGAGCCCGCGGGGA-GCGCGCGCAATTGTCGCGCCCGG 59  
Db |||||  
Qy 4 ACCGCTGTCGCGCGGCTTGAGCCCGCGGGAGCGCGCAATTGTCGCGCCCGG 63  
Db |||||  
Qy 60 GGGGCGGCTCCCGGCATCTTCGGCGGACCAAGACTACAGAGGGGCGGCTGG 119  
Db |||||  
Qy 64 GGGGCGGCTCCCGGCATCTTCGGCGGACCAAGACTACAGAGGGGCGGCTGG 123  
Db |||||  
Qy 120 GATGCGGCTCGCGGCGGCGGAGTACAAAGCGGCGACCTGCTTCGCCAAGATGAA 179  
Db |||||  
Qy 124 GATGCGGCTCGCGGCGGCGGAGTACAAAGCGGCGACCTGCTTCGCCAAGATGAA 183  
Db |||||  
Qy 180 GGGCTACCGCACTGGCGGCGGCGGATGATGAATCCAGAGGCGCTGTGAAGCTCC 239  
Db |||||  
Qy 184 GGGCTACCGCACTGGCGGCGGCGGATGATGAATCCAGAGGCGCTGTGAAGCTCC 243  
Db |||||  
Qy 240 AGCAACAAGTATCTATCTCTTTTGGCACCCATGCAATCTTAGTGTCCCA 299  
Db |||||  
Qy 244 AGCAACAAGTATCTATCTCTTTTGGCACCCATGCAATCTTAGTGTCCCA 303  
Db |||||  
Qy 300 AGACTTTTCCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAACAAACGAAGG 359  
Db |||||  
Qy 304 AGACTTTTCCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAACAAACGAAGG 363  
Db |||||  
Qy 360 ATTTAACAAGAGTGTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGCTACCA 419  
Db |||||  
Qy 364 ATTTAACAAGAGTGTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGCTACCA 423  
Db |||||  
Qy 420 GGCATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTTGAATCTCGAGATGC 479  
Db |||||  
Qy 424 GGCATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTTGAATCTCGAGATGC 483  
Db |||||  
Qy 480 AAGCAGTTCAGCAACAGAGTGTAGAGTAGAAGAGTGAAGGCAAGAAAGATGA 539  
Db |||||  
Qy 484 AAGCAGTTCAGCAACAGAGTGTAGAGTAGAAGAGTGAAGGCAAGAAAGATGA 543  
Db |||||  
Qy 540 AAAAGCAGGTCAAAACGGAAGAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCCG 599  
Db |||||

Db 544 AAAAGCAGGTCAAAACGGAAGAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCCG 603  
Qy |||||  
Db 600 GAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAAAGAGAGGAAACAAACAGCTC 659  
Qy |||||  
Db 604 GAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAAAGAGAGGAAACAAACAGCTC 663  
Qy |||||  
Db 660 TGAGGCTGAGATGCGGGCAACGACACAAAGAAACACAACTTCAGACTTGCAGAAACCCAG 719  
Qy |||||  
Db 664 TGAGGCTGAGATGCGGGCAACGACACAAAGAAACACAACTTCAGACTTGCAGAAACCCAG 723  
Qy |||||  
Db 720 TGAAGGACCTTAACCTACCATTAATGAATGCTGCATTAATTAAGAGAAACCAAGAGGTTAT 779  
Qy |||||  
Db 724 TGAAGGACCTTAACCTACCATTAATGAATGCTGCATTAATTAAGAGAAACCAAGAGGTTAT 783  
Qy |||||  
Db 780 ATGTTTGGTGTCTAATATTTCTTGGATTGATATGAACCAACACATAGTCTCTTGTGCA 839  
Qy |||||  
Db 784 ATGTTTGGTGTCTAATATTTCTTGGATTGATATGAACCAACACATAGTCTCTTGTGCA 843  
Qy |||||  
Db 840 TTGACAGAACCCAGTTTGTATGTATATTCATATTCCTCTCTCTGTGTGTTCGGGG 899  
Qy |||||  
Db 844 TTGACAGAACCCAGTTTGTATGTATATTCATATTCCTCTCTCTGTGTGTTCGGGG 903  
Qy |||||  
Db 900 GAAAAGACATTTAGCCCTTTTAAAGTACGATTAATTCATGTTATTCGTTGCA 959  
Qy |||||  
Db 904 GAAAAGACATTTAGCCCTTTTAAAGTACGATTAATTCATGTTATTCGTTGCA 963  
Qy |||||  
Db 960 TGAAGTTGCCCTTAACCACTAAGGATTAATCAAGATTTTTCGGGACGCTTATACATGCTA 1019  
Qy |||||  
Db 964 TGAAGTTGCCCTTAACCACTAAGGATTAATCAAGATTTTTCGGGACGCTTATACATGCTA 1023  
Qy |||||  
Db 1020 GGATC 1024  
Qy |||||  
Db 1024 GGATC 1028  
RESULT 2  
AF151900 1830 bp mRNA linear PRI 18-MAY-2000  
LOCUS Homo sapiens CGI-142 protein mRNA, complete cds.  
DEFINITION  
ACCESSION AF151900  
VERSION AF151900.1 GI:4929752  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 1830)  
Lai, C.H., Chou, C.Y., Ch'ang, L.Y., Liu, C.S. and Lin, W.  
Identification of novel human genes evolutionarily conserved in  
Caenorhabditis elegans by comparative proteomics  
Genome Res. 10 (5), 703-713 (2000)  
JOURNAL 20272150  
MEDLINE 10810093  
PUBMED  
REFERENCE 2 (bases 1 to 1830)  
Lin, W.-C.  
AUTHORS  
TITLE Direct Submission  
JOURNAL Submitted (17-MAY-1999) Institute of Biomedical Sciences, Academia  
Sinica, No. 128, Sec. II, Academia Road, Taipei 115, Taiwan  
FEATURES  
Location/Qualifiers  
1. .1830  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
125. .736  
/codon\_start=1  
/product="CGI-142 protein"  
/protein\_id="AAD34137.1"  
/db\_xref="GI:4929753"  
/translation="MARPRREYKAGDLVFAVMKGYPHWPARIDELPEGAVKPPANKY  
PIFFGTHETAFGLPKDLFPYKEYDKFGKSNKRKFNGLWEIENNPVGFYQAI  
QQSSSTEGEGNTADASSEEDGVEEDGKGRKNEKAGSKKSYTSKSSKQSR  
KSPGDEDDKDCKEENKSSSEGGDAGNDTRNTTSDLQKTSST"  
source  
CDS

BASE COUNT	588 a	342 c	385 g	515 t	
ORIGIN					
Query Match	98.9%; Score 1013; DB 9; Length 1830;				
Best Local Similarity	99.9%; Pred. No. 1.1e-237;				
Matches 1024; Conservative	0; Mismatches 0; Indels 1; Gaps 1;				
QY	1	ACCGCTCGTCCGCGCGCTTTGAGCGCCGCGGGGAGCGG-CGCAATTCGTGCGCCCGCGG	59		
Db	4	ACCGCTCGTCCGCGCGCTTTGAGCGCCGCGGGAGCGCGCAATTCGTGCGCCCGCGG	63		
QY	60	GGGGGGCGCTCCCGCGCATCTTCGCGCGGACCAAGGACTACCAAGAGGGAGGGCGCTGG	119		
Db	64	GGGGGGCGCTCCCGCGCATCTTCGCGCGGACCAAGGACTACCAAGAGGGAGGGCGCTGG	123		
QY	120	GATGGCGCGTCCGCGCGCCCGCGAGTACAAAGCGGGCGACTGGTCTTCGCCAAGATGAA	179		
Db	124	GATGGCGCGTCCGCGCGCCCGCGAGTACAAAGCGGGCGACTGGTCTTCGCCAAGATGAA	183		
QY	180	GGGCTACCCGCACTGGCGCGCCCGCGGATTTGATGAATCCCAAGAGGGCGCTGTGAAGCTCC	239		
Db	184	GGGCTACCCGCACTGGCGCGCCCGCGGATTTGATGAATCCCAAGAGGGCGCTGTGAAGCTCC	243		
QY	240	AGCAACAGTATCCTATCTCTTTTGGCACCATGAACTGCAATTTCTAGGTCCCAA	299		
Db	244	AGCAACAGTATCCTATCTCTTTTGGCACCATGAACTGCAATTTCTAGGTCCCAA	303		
QY	300	AGACCTTTTCCATATAAGGAGTACAAGACAAGTTTGGAAAGTCAAAACAACGGAAGG	359		
Db	304	AGACCTTTTCCATATAAGGAGTACAAGACAAGTTTGGAAAGTCAAAACAACGGAAGG	363		
QY	360	ATTAAACGAAGGATTTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACCA	419		
Db	364	ATTAAACGAAGGATTTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACCA	423		
QY	420	GGCAATTCAGCAACAGAGCTTTCAGAACTGAGGAGAGAGGTGGAAATCTCGAGATGC	479		
Db	424	GGCAATTCAGCAACAGAGCTTTCAGAACTGAGGAGAGAGGTGGAAATCTCGAGATGC	483		
QY	480	AACGAGTGAGGAAGAGTGTATAGATGAGAGAGATGGAAGAGGCAAAAGAAAGATGA	539		
Db	484	AACGAGTGAGGAAGAGTGTATAGATGAGAGAGATGGAAGAGGCAAAAGAAAGATGA	543		
QY	540	AAAAGCAGGCTCAAAACGGAAGTATCATATCTTCAAAGAAATCTCTTAAACAGTCCCG	599		
Db	544	AAAAGCAGGCTCAAAACGGAAGTATCATATCTTCAAAGAAATCTCTTAAACAGTCCCG	603		
QY	600	GAATCTCCAGGAGATGAAGATGACAAGACTGCAAGAGAGAGAAACAAAAGCAGCTC	659		
Db	604	GAATCTCCAGGAGATGAAGATGACAAGACTGCAAGAGAGAGAAACAAAAGCAGCTC	663		
QY	660	TGAGGTTGAGATGCGGGCAACGACACAAGAAACAACTTCAGACTTCGAGAAACCAG	719		
Db	664	TGAGGTTGAGATGCGGGCAACGACACAAGAAACAACTTCAGACTTCGAGAAACCAG	723		
QY	720	TGAAGGACCTAACTACATAATGAATGCTGTCATATTAAGAGAAACCAACAAGAGTTAT	779		
Db	724	TGAAGGACCTAACTACATAATGAATGCTGTCATATTAAGAGAAACCAACAAGAGTTAT	783		
QY	780	ATGTTTGGTTGCTAATATTTCTGGATTGATATGAACCAACACATAGTCTTTGTTGTC	839		
Db	784	ATGTTTGGTTGCTAATATTTCTGGATTGATATGAACCAACACATAGTCTTTGTTGTC	843		
QY	840	TTACAGAACCCAGTTTGTATGATATTCATATTCCTCTGTTGTTTCCGGGG	899		
Db	844	TTACAGAACCCAGTTTGTATGATATTCATATTCCTCTGTTGTTTCCGGGG	903		
QY	900	GAAGAAGACATTTTAGCCCTTTTAAAGTTACTGATTAATTTCACTGTTATTTGGTTGCA	959		
Db	904	GAAGAAGACATTTTAGCCCTTTTAAAGTTACTGATTAATTTCACTGTTATTTGGTTGCA	963		
QY	960	TGAAGTTGCCCTTAAACCACTAAGGATTAACAAGATTTTCCGAGACTTATACATGTCTA	1019		
Db	964	TGAAGTTGCCCTTAAACCACTAAGGATTAACAAGATTTTCCGAGACTTATACATGTCTA	1023		
QY	1020	GGATC 1024			
Db	1024	GGATC 1028			
RESULT 3					
BD156109	1973 bp DNA linear PAT 17-JAN-2003				
LOCUS	Primer for synthesizing full-length cDNA and use thereof.				
DEFINITION	BD156109.1 GI:27861867				
ACCESSION	BD156109.1				
VERSION	JP 2002191363-A/10952.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1973)				
AUTHORS	Ota, T., Isogai, T., Nishikawa, T., Hayaashi, K., Saito, K., Yamamoto, J., Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.				
TITLE	Primer for synthesizing full-length cDNA and use thereof				
JOURNAL	Patent: JP 2002191363-A 10952 09-JUL-2002;				
COMMENT	HELIX RESEARCH INSTITUTE OS Homo sapiens (human) PN JP 2002191363-A/10952 PD 09-JUL-2002 PF 28-JUL-2000 JP 2000280990 PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU, PI KEIICHI NAGAI, TETSUJI OTSUKI PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof PH Key Location/Qualifiers (156)..(764).				
FT	CDS	Location/Qualifiers			
source	1..1973	/organism="Homo sapiens"			
		/mol_type="genomic DNA"			
		/db_xref="taxon:9606"			
BASE COUNT	627 a	367 c	420 g	559 t	
ORIGIN					
Query Match	98.8%; Score 1012; DB 6; Length 1973;				
Best Local Similarity	99.9%; Pred. No. 2e-237;				
Matches 1023; Conservative	0; Mismatches 0; Indels 1; Gaps 1;				
QY	2	CGCGTCTGTCGCGCGCGCTTGAGCGCCGCGGGAGCGG-CGGCAATTCGTGCGCCCGCGG	60		
Db	36	CGCGTCTGTCGCGCGCGCTTGAGCGCCGCGGGAGCGCGCGCAATTCGTGCGCCCGCGG	95		
QY	61	GGGGCGGCTCCCGCGCATCTTCGCGCGACCAAGGACTACCAAGAGGGAGGGCGGTGGG	120		
Db	96	GGGGCGGCTCCCGCGCATCTTCGCGCGACCAAGGACTACCAAGAGGGAGGGCGGTGGG	155		
QY	121	ATGGCGGCTCCCGCGCGCCCGCGAGTACAAAGCGGGCGACCTGGTCTTCGCCAAGATGAAG	180		
Db	156	ATGGCGGCTCCCGCGCGCCCGCGAGTACAAAGCGGGCGACCTGGTCTTCGCCAAGATGAAG	215		
QY	181	GGCTACCCGCACTGGCGCGCCCGGATTTGATGAATCCCAAGAGGGCGCTGTGAAGCTCCA	240		
Db	216	GGCTACCCGCACTGGCGCGCCCGGATTTGATGAATCCCAAGAGGGCGCTGTGAAGCTCCA	275		
QY	241	GCAACAAGTATCTTATCTTTTGGCACCCATGAACTGCAATTTCTAGGTCCCAA	300		
Db	276	GCAACAAGTATCTTATCTTTTGGCACCCATGAACTGCAATTTCTAGGTCCCAA	335		
QY	301	GACCTTTTCCATATAAGGAGTACAAAGAGTGTGGAAAGTCAAAACAACGGAAGGA	360		

```
Db 336 GACCTTTTCCATATAAGAGGAGTACAAAGACAAAGTTTGGAAAGCTCAAAACAAACGGAAGGA 395
Qy 361 TTTACGAGGATTCGGGAATAGAAATATACCCAGGAGTAAAGTTTACTTGGCTACACG 420
Db 396 TTTACGAGGATTCGGGAATAGAAATATACCCAGGAGTAAAGTTTACTTGGCTACACG 455
Qy 421 GCAATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGAGTGGAAATCTGCAGATGCA 480
Db 456 GCAATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGAGTGGAAATCTGCAGATGCA 515
Qy 481 AGCAGTGAAGAAAGAGTGTATAGATAGAGTGAAGAGTGAAGAGGCAAAAGAAAGAAATGAA 540
Db 516 AGCAGTGAAGAAAGAGTGTATAGATAGAGTGAAGAGTGAAGAGGCAAAAGAAAGAAATGAA 575
Qy 541 AAAGCAGGCTCAAAACGGAAGAGTCATATCTTCAAGAAATCTCTTAAACAGTCCCGG 600
Db 576 AAAGCAGGCTCAAAACGGAAGAGTCATATCTTCAAGAAATCTCTTAAACAGTCCCGG 635
Qy 601 AAATCTCCAGGAGATGAAGATGACAAAGACTCAAAAGAGAGAGGAGGAGGAGGAGGAGTCT 660
Db 636 AAATCTCCAGGAGATGAAGATGACAAAGACTCAAAAGAGAGAGGAGGAGGAGGAGTCT 695
Qy 661 GAGGTGAGATGCGGGCAACGACACAGAAACACAACTTCAGACTTGCAGAAAACCCAGT 720
Db 696 GAGGTGAGATGCGGGCAACGACACAGAAACACAACTTCAGACTTGCAGAAAACCCAGT 755
Qy 721 GAAGGACCTTAACCTAATGATGCTGCATATTAAGAGAAACCAAGAGGTTATA 780
Db 756 GAAGGACCTTAACCTAATGATGCTGCATATTAAGAGAAACCAAGAGGTTATA 815
Qy 781 TGTTTGGTGTCTAATATCTTGGATTTGATATGAACCAACATAGTCTCTTGTGTCTAT 840
Db 816 TGTTTGGTGTCTAATATCTTGGATTTGATATGAACCAACATAGTCTCTTGTGTCTAT 875
Qy 841 TGACAGAACCCAGTTGTATGATCATTTATCATATTCCTCTCTGTGTGTGTTCGGGGG 900
Db 876 TGACAGAACCCAGTTGTATGATCATTTATCATATTCCTCTCTGTGTGTGTTCGGGGG 935
Qy 901 AAAACACATTTAGCTTTTAAAGTTACTGATTTAATTCATGTTTATTTGGTTGAT 960
Db 936 AAAACACATTTAGCTTTTAAAGTTACTGATTTAATTCATGTTTATTTGGTTGAT 995
Qy 961 GAAGTTGCTTAAACCTAAGGATTAACAGATTTTTCGAGACTTATATCATGCTAG 1020
Db 996 GAAGTTGCTTAAACCTAAGGATTAACAGATTTTTCGAGACTTATATCATGCTAG 1055
Qy 1021 GATC 1024
Db 1056 GATC 1059
```

```
RESULT 4
AK001280
LOCUS
DEFINITION
  Homo sapiens cDNA FLJ10418 fis, clone NT2RP1000130, moderately
  similar to HEPATOMA-DERIVED GROWTH FACTOR.
ACCESSION
  AK001280
VERSION
  AK001280.1 GI:7022435
KEYWORDS
  oligo capping; fis (full insert sequence).
SOURCE
  Homo sapiens
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1
  Isogai,T., Ota,T., Hayaashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
  Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
  Wagatsuma,M., Hosoi,T., Kaku,Y., Kodaira,H., Kondo,H.,
  Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
  Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
  Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
  Nakamura,Y., Nagahara,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
  NEDO human cDNA sequencing project
  JOURNAL
  Unpublished
```

```
REFERENCE
  2 (bases 1 to 1973)
  Isogai,T. and Otsuki,T.
  Direct Submission
  Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,
  Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
  (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
  NEDO human cDNA sequencing project supported by Ministry of
  International Trade and Industry of Japan; cDNA full insert
  sequencing; Research Association for Biotechnology; cDNA library
  construction, 5'- & 3'-end one pass sequencing and clone selection;
  Helix Research Institute (supported by Japan Key Technology Center
  etc.) and Department of Virology, Institute of Medical Science,
  University of Tokyo.
FEATURES
  Location/Qualifiers
    1..1973
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="NT2RP1000130"
    /cell_line="NT2"
    /cell_type="teratocarcinoma"
    /clone_lib="NT2RP1"
    /note="cloning vector: pUC19FL3-mRNA from NT2 neuronal
    precursor cells after 48-hours retinoic acid (RA)
    induction."
    156..767
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="BAA91597.1"
    /db_xref="GI:7022436"
    /translation="MARPRPREYKAGDLVFAKMGYPHWPARIDELPEGAVKPPANKY
    PIPFGTHETAFLGKDLFPYKEYKDFKSNKRGNEGLWEIENNVKVTGYQAI
    QQSSSTEGEGNTADASSEEDRVEDGKGRKKEKAGSKKSKTSKSSKQSR
    KSPDEDDKCKEENKSSSEGGAGNDTRNTTSDLOKTSBGT"
BASE COUNT
  627 a 367 c 420 g 559 t
```

```
Query Match 98.8%; Score 1012; DB 9; Length 1973;
Best Local Similarity 99.9%; Pred. No. 2e-237;
Matches 1023; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 2 CCGCTCGTCCCGCGGCTTGAGGCGCGGCGGAGCGC-GCGCAATTCGTCCGCGCGCGG 60
Db 36 CCGCTCGTCCCGCGGCTTGAGGCGCGGCGGAGCGCGGCGCAATTCGTCCGCGCGCGG 95
Qy 61 GGGGCGGCTCCCGGCATCTTCGCGCGACCAAGACTACCAAGAGGGAGCGGCTGGG 120
Db 96 GGGGCGGCTCCCGGCATCTTCGCGCGACCAAGACTACCAAGAGGGAGCGGCTGGG 155
Qy 121 ATGGCGGCTCCCGGCGCGCGAGTACAAAGCGGCGACCTGCTTCGCGCAAGATGAAG 180
Db 156 ATGGCGGCTCCCGGCGCGCGAGTACAAAGCGGCGACCTGCTTCGCGCAAGATGAAG 215
Qy 181 GGCTACCGCACTGGCGCGCGCGGATGATGAACCTCCAGAGGGCGGCTGTGAAGCTCCA 240
Db 216 GGCTACCGCACTGGCGCGCGCGGATGATGAACCTCCAGAGGGCGGCTGTGAAGCTCCA 275
Qy 241 GCAAAACAGTATCTTCTTTTGGGACCAACATGAACATGCAATTTCTAGGTCCCAA 300
Db 276 GCAAAACAGTATCTTCTTTTGGGACCAACATGAACATGCAATTTCTAGGTCCCAA 335
Qy 301 GACCTTTTCCATATAGGAGTACAAAGAGTTTGGAAAGTCAAAACAAACGGAAGGA 360
Db 336 GACCTTTTCCATATAGGAGTACAAAGAGTTTGGAAAGTCAAAACAAACGGAAGGA 395
Qy 361 TTTAACAGAGGATTTGGGAAATAGAAATTAACCCAGAGTAAAGTTTACTTGGCTACACG 420
Db 396 TTTAACAGAGGATTTGGGAAATAGAAATTAACCCAGAGTAAAGTTTACTTGGCTACACG 455
Qy 421 GCAATTCAGCAACAGAGCTCTTCAAGAACTGAGGAGAGAGTGGAAATCTGCAGATGCA 480
Db 456 GCAATTCAGCAACAGAGCTCTTCAAGAACTGAGGAGAGAGTGGAAATCTGCAGATGCA 515
```



```
Qy 481 AGCAGTGAAGGAGGTGATAGAGTGAAGAGATGAAAGGCAAAAGAAAGAAATGAA 540
Db 516 AGCAGTGAAGGAGGTGATAGAGTGAAGAGATGAAAGGCAAAAGAAAGAAATGAA 575
Qy 541 AAACGAGGCTCAAAACGGAAAAGTCAATATCTTCAAGAAATCTCTAAACAGTCCCGG 600
Db 576 AAACGAGGCTCAAAACGGAAAAGTCAATATCTTCAAGAAATCTCTAAACAGTCCCGG 635
Qy 601 AAATCTCCAGAGATGAAGATGACAAAGACTGCAAGAGAGGAAAACAAAAGCAGTCT 660
Db 636 AAATCTCCAGAGATGAAGATGACAAAGACTGCAAGAGAGGAAAACAAAAGCAGTCT 695
Qy 661 GAGGCTGAGATGCGGCAACGACACAAAGAAACAACTTCAGACTTGCAGAAAACCACT 720
Db 696 GAGGCTGAGATGCGGCAACGACACAAAGAAACAACTTCAGACTTGCAGAAAACCACT 755
Qy 721 GAAGGACCTTAACCTACCATATGATGCTGCATATTAAGAGAACCAAGAGTTTATA 780
Db 756 GAAGGACCTTAACCTACCATATGATGCTGCATATTAAGAGAACCAAGAGTTTATA 815
Qy 781 TGTGTGGTGTCTAATATCTTGGATTTGATATGAACCAACACATAGTCTCTTGTGTCT 840
Db 816 TGTGTGGTGTCTAATATCTTGGATTTGATATGAACCAACACATAGTCTCTTGTGTCT 875
Qy 841 TGACAGAACCCAGTTTGTATGATATTAATTCATATTCCTCTCTGTGTGTTCGCGGGG 900
Db 876 TGACAGAACCCAGTTTGTATGATATTAATTCATATTCCTCTCTGTGTGTTCGCGGGG 935
Qy 901 AAAAGACATTTAGCCTTTTAAAGTACTGATTTAAATTCATTTATTTGTTGTCAT 960
Db 936 AAAAGACATTTAGCCTTTTAAAGTACTGATTTAAATTCATTTATTTGTTGTCAT 995
Qy 961 GAAGTGGCCCTTAACCACTAAGGATTTCAAGATTTTGGCAGACTTATACATGCTAG 1020
Db 996 GAAGTGGCCCTTAACCACTAAGGATTTCAAGATTTTGGCAGACTTATACATGCTAG 1055
Qy 1021 GATC 1024
Db 1056 GATC 1059

RESULT 5
AB029156 1427 bp mRNA linear PRI 19-FEB-2000
LOCUS Homo sapiens HRP-3 mRNA, complete cds.
DEFINITION AB029156
ACCESSION AB029156
VERSION AB029156.1 GI:6855467
KEYWORDS HRP-3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
Ikegane, K., Yamamoto, M., Kishima, Y., Enomoto, H., Yoshida, K.,
Suemura, M., Kishimoto, T. and Nakamura, H.
A new member of a hepatoma-derived growth factor gene family can
translocate to the nucleus
Biochem. Biophys. Res. Commun. 266 (1), 81-87 (1999)
JOURNAL 20050055
MEDLINE 10581169
PUBMED 10581169
REFERENCE 2 (bases 1 to 1427)
Ikegane, K. and Nakamura, H.
Direct Submision
AUTHORS Submitted (21-JUN-1999) Kazuhiro Ikegane, Osaka University Medical
SCHOOL, Department of Molecular Medicine; 2-2 Yamada-oka, Suita,
Osaka 565-0871, Japan (E-mail: nakamura@med3.med.osaka-u.ac.jp,
Tel:81-6-6879-3837(ex.3837), Fax:81-6-6879-3839)
JOURNAL Location/Qualifiers
FEATURES
source
1..1427
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="15"
```

```

/map="15q25"
/cell_type="lymphocytes"
1..1427
/gene="HRP-3"
/genes="HRP-3"
121..732
/codon_start=1
/product="HRP-3"
/db_xref="GI:6855468"
/protein_id="BAA90477.1"
/translation="MARPRPREYKAGDLVPAKWKGYPHHPARIDELPEGAVKPKPANKY
PIFFGTHTAFLGPKDADLPFYKDFKFSKRNKRGFNGLMEIENNPGVKFYGYQAI
QQSSSETTEGGNTADLSEEDRVEDGKRGKXKRGKRSYTSKSSKQSR
KSPGDBDDDKCKEENKSSSEGGDNGDNRNTTSDLOKTSEGT"
BASE COUNT 454 a 287 c 327 g 359 t
ORIGIN
Query Match 97.8%; Score 1001; DB 9; Length 1427;
Best Local Similarity 99.8%; Pred. No. 9.4e-235;
Matches 1023; Conservative 0; Mismatches 0; Indels 2; Gaps 2;
Qy 1 ACCGCTCGTCCGCCCGCTTGAGGCCCGCGGGAGCGC-GCGCAATTCTGTCGCCCGCGG 59
Db 1 ACCGCTCGTCCG-CCGGCTTGAGGCCCGCGGGAGCGCGCAATTCTGTCGCCCGCGG 59
Qy 60 GGGGGCGGCTCTCCCGCATCTTCGCGCGCAACCAAGGACTACCAAGAGGAGGAGCGGCTGG 119
Db 60 GGGGGCGGCTCTCCCGCATCTTCGCGCGCAACCAAGGACTACCAAGAGGAGGAGCGGCTGG 119
Qy 120 GATGGCGCTGCGCGCCCGCGAGTACAAAGCGGCGGACCTGGTCTTTCGCCAAGATGAA 179
Db 120 GATGGCGCTGCGCGCCCGCGAGTACAAAGCGGCGGACCTGGTCTTTCGCCAAGATGAA 179
Qy 180 GGGCTACCGCATCTGCGCGCCCGCGGATTTGATGAATCTCCAGAGGCGCTGTGAAGCTCC 239
Db 180 GGGCTACCGCATCTGCGCGCCCGCGGATTTGATGAATCTCCAGAGGCGCTGTGAAGCTCC 239
Qy 240 AGCAAAACAAGTATCTCTCTTTTGGCACCCCATGAAACTGCACTTTCTAGTCTCCAA 299
Db 240 AGCAAAACAAGTATCTCTCTTTTGGCACCCCATGAAACTGCACTTTCTAGTCTCCAA 299
Qy 300 AGACCTTTTCCATATAAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACGGAAGG 359
Db 300 AGACCTTTTCCATATAAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACGGAAGG 359
Qy 360 ATTTAAGGAAGGATTTGTGGAAATAGAAAATACCAGGATTAAGTTTACTGGCTACCA 419
Db 360 ATTTAAGGAAGGATTTGTGGAAATAGAAAATACCAGGATTAAGTTTACTGGCTACCA 419
Qy 420 GGCAATTCAGCAACAGAGCTCTTCAGAAAAGTGAAGGAGAGGTGGAAAATACTGCAGATGC 479
Db 420 GGCAATTCAGCAACAGAGCTCTTCAGAAAAGTGAAGGAGAGGTGGAAAATACTGCAGATGC 479
Qy 480 AAGCAGTGAAGGAAGAGGTGATAGAGTAGAAGAGATGGAAGAAAGGCAAAAGAAAGATGA 539
Db 480 AAGCAGTGAAGGAAGAGGTGATAGAGTAGAAGAGATGGAAGAAAGGCAAAAGAAAGATGA 539
Qy 540 AAAAGCAGGCTCAAAACGGAAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCCG 599
Db 540 AAAAGCAGGCTCAAAACGGAAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCCG 599
Qy 600 GAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAGAGAGGAAAACAAAAGCAGCTC 659
Db 600 GAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAGAGAGGAAAACAAAAGCAGCTC 659
Qy 660 TGAAGGCTGAGATGCGGCAACGACACAAAGAAACAACTTCAGACTTGCAGAAAACCCAG 719
Db 660 TGAAGGCTGAGATGCGGCAACGACACAAAGAAACAACTTCAGACTTGCAGAAAACCCAG 719
Qy 720 TGAAGGACCTAATACCATTAATGAATGCTGCATATTAAGAGAAAACCAAGAGGTTAT 779
Db 720 TGAAGGACCTAATACCATTAATGAATGCTGCATATTAAGAGAAAACCAAGAGGTTAT 779
```





Db	4	CCGCTCGAGCGCCGCGCTCGAGGCCCGGGAGCGCGCGGTTCGTTCGTCCCGCGGA	63
Qy	61	GGGCGGCGCTCCGCGCATCTTCGCGGCGGACCAAGGACTTACAGGAAGGGAGCGGCTGGG	120
Db	64	GGGCGGCGCTCCGCGCATCTTCGCGGCGGACCAAGGACAGCGGCGAGGGGAAACCAACGGG	123
Qy	121	ATGGCGGCGTCCGCGGCCCGCGAGTACAAAGCGGCGACCTGGTCTTCGCGCAAGATGAAG	180
Db	124	ATGGCGGCGTCCGCGGCCCGCGAGTACAAAGCGGCGACCTGGTCTTCGCGCAAAATGAAG	183
Qy	181	GGCTACCGCGACTGGCGGCCCGCGATTGATGAATCCCGAGAGGCGCTGTGAAGCCCTCCA	240
Db	184	GGCTACCGCGACTGGCGGCCCGCGATTGATGAATCCCGAGAGGCGCGTGAAGCCTCCA	243
Qy	241	GCAAAACAAGTATCCTATCTCTTTTGGCAACCCATCAAACTGCAATTCCTAGGTCCTAAA	300
Db	244	GCAAAACAAGTATCCTATCTCTTTTGGCAACCCATCAAACTGCAATTCCTAGGTCCTAAA	303
Qy	301	GACCTTTTCCATATAAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACAACGGAAGGA	360
Db	304	GACCTTTTCCATATAAGGAAATACAAAGACAAGTTTGGAAAGTCAAAACAACGGAAGGA	363
Qy	361	TTTAAACGAAGGATTGTGGGAAATAGAAAATAAACCCAGGAGTAAAGTTTACTTGGCTACCA	420
Db	364	TTTAAACGAAGGATTGTGGGAAATAGAAAATAAACCCAGGAGTAAAGTTTACTTGGCTACCA	423
Qy	421	GCAATTCAGCAACAGAGCTCTTCAGAACTGAGGGGAGAGGTGGAATATCTCGCATGCA	480
Db	424	ACAATTCAGCAACAGAGCTCTTCAGAACTGAGGGGAGAGGTGGAATATCTCGCATGCA	483
Qy	481	AGCAGTCAGGAAGAAGGTGATAGAGTAGAAGATGGAAGAAGTCAAAAGGCAAAAGAAATGAA	540
Db	484	AGCAGTCAGGAAGAAGGTGACAGAGT--AGAAGTGGAAAGGCAAGAGAAAGATGAA	540
Qy	541	AAAGCGGCTCAAAACGGGAAAGAGTCAATATCTTCAAAGAAATCTCTTAAACAGTCCCGG	600
Db	541	AAAGCGGCTCAAAACGGGAAAGAGTCTTCAACATCTTCAAAGAAATCTCTTAAACAGTCCCGG	600
Qy	601	AAATCTCCAGGAGTGAAGATGCACAAAGACTGCAAGAGGAGGAAACAAACGACGCTCT	660
Db	601	AAATCTCCAGGAGTGAAGATGCACAAAGACTGCAAGAGGAGGAGGAAACAAACGACGCTCT	660
Qy	661	GAGGTCGAGATGCGGGCAACGACACAAGAAACAACTTTCAGACTTGCAGAAAACCGAGT	720
Db	661	GAGGTCGAGATGCTGGCAATGCACAAAGAAACACGACTTTCAGACTTGCAGAAAACCGAGT	720
Qy	721	GAGGGACCTTAACCTACCATATGATGCTGCAATTAAGAGAAACCAACAAGAGTTATA	780
Db	721	GAGGGACCTTAACCTACCGTAATGATGCTGCAATTAAGAGAAACCAACAAGAGTTATA	780
Qy	781	TGTTTGTGTGCTAATATCTTGGATTGATATGAACCAACACATAGTCTCTTGTGTGCTAT	840
Db	781	TGTTTGTGTGCTGA---TTCTGGATTGATATGAACCAAC---AGTCTTGTGTGTGCTAT	833
Qy	841	TGACAGAACCCCGAGTTTGTATGTACATTAATTCATTAATTCATGTTATTCGTTGGGGGG	900
Db	834	TGACAAAGCCCGAGTGTGTGCTAAT---ACATTCCTCTCTGTTCTCTGTAAGGGAA	889
Qy	901	AAAAGACATTTTAGCCTTTTAAAGTACTGATTTAAATTCATGTTATTCGTTGTCAT	960
Db	890	AAAAGACATTTTAAACCTTTTAAAGGTTATGACTTAATTTATGTTATTCGTTGTCAT	949
Qy	961	GAGTTCGCCCTTAACCACTAAGGATTATCAAGATTTTTCGCGACACTTATACATGCTAG	1020
Db	950	GAGTTCGCCCTTAACCACTAAGATGATCAAGATTTT---GCAGACTTCTCCGTGCTAG	1006
Qy	1021	GATC 1024	
Db	1007	GCTC 1010	

RESULT 12  
BD147622

LOCUS	B0147622		737 bp	DNA	linear	PAT 17-JAN-2000
DEFINITION	Primer for synthesizing full-length cDNA and use thereof.					
ACCESSION	B0147622					
VERSION	B0147622.1	GI:27853380				
KEYWORDS	JP 2002191363-A/2465					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 737)					
AUTHORS	Ota,T., Isogai,T., Nishikawa,T., Hayaashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.					
TITLE	Primer for synthesizing full-length cDNA and use thereof					
JOURNAL	Patent: JP 2002191363-A/2465 09-JUL-2002; HELIIX RESEARCH INSTITUTE					
COMMENT	OS Homo sapiens (human) PN JP 2002191363-A/2465 PD 09-JUL-2002 PF 28-JUL-2000 JP 2000280990 PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU SAITO, PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU, PI KEIICHI NAGAI,TETSUI OTSUKI PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC 10, CC C12P21/02,C12P1/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key					
FT source	Location/Qualifiers					
FT	1..737 /organism='Homo sapiens (human)'. 1..737 /organism='Homo sapiens' /mol_type='genomic DNA' /db_xref='taxon:9606'					
BASE COUNT	227 a 180 c 204 g 117 t					9 others
ORIGIN						
Query Match	60.9%; Score 623.6;					DB 6; Length 737;
Best Local Similarity	96.2%; Pred. No. 3.8e-142;					
Matches	676; Conservative 0; Mismatches 23; Indels 4; Gaps 4;					
Qy	2	CCGCTCGTCGCCGGCCTTGAGGCCTTCCGCGGAGCGC-GGCAATTCTGCGCCCGCGG 60				
Db	36	CCGCTCGTCGCCGGCCTTGAGGCCTTCCGCGGAGCGCCTTCTGCGCCCGCGG 95				
Qy	61	GGGCGCGCTCCCGCATCTTCGCGCGACTACCAGGAAGGAGCGGTGGG 120				
Db	96	GGGCGCGCTCCCGCATCTTCGCGCGACTACCAGGAAGGAGCGGTGGG 155				
Qy	121	ATGCGCGTCCGCGGCCCGGAGTAGTAACAAGCGGCGACTGTCTTCGCGAAGTAAG 180				
Db	156	ATGCGCGTCCGCGGCCCGGAGTAGTAACAAGCGGCGACTGTCTTCGCGAAGTAAG 215				
Qy	181	GGCTACCGCACATGGCGGCCGCGGTTGATGAATCCCAGAGGCGCTGGAAGCCTCCA 240				
Db	216	GGCTACCGCACATGGCGGCCGCGGTTGATGAATCCCAGAGGCGCTGGAAGCCTCCA 275				
Qy	241	GCAAAACAAGTATCCTATCTTTTGGCACCCTATGAACTGCAATTCCTAGGTCCCAA 300				
Db	276	GCAAAACAAGTATCCTATCTTTTGGCACCCTATGAACTGCAATTCCTAGGTCCCAA 335				
Qy	301	GACCTTTTCCATATAAGGAGTCAAGAACAAGTTGGAAAGTCAACAACGGAAGGA 360				
Db	336	GACCTTTTCCATATAAGGAGTCAAGAACAAGTTGGAAAGTCAACAACGGAAGGA 395				
Qy	361	TTTAAACAAGGATGTGGGAAATAGAAAATAACCCAGGAGTAAAGTTACTGCTACCAG 420				
Db	396	TTTAAACAAGGATGTGGGAAATAGAAAATAACCCAGGAGTAAAGTTACTGCTACCAG 455				
Qy	421	GCAATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTCGGAATATCTGCAGATGCA 480				







PT version, useful e.g. in study and development of drugs for hepatoma -  
XX  
PS Claim 1; Page 14; 27pp; Chinese.

This sequence represents the cDNA sequence encoding the human hepatoma-derived growth factor (HDGF2) protein. The HDGF2 sequence shows considerable homology to the type I version. The protein has cytostatic activity. The HDGF2 protein, derivatives and polynucleotides are useful e.g. in the study and development of drugs for the treatment of hepatoma.

SQ Sequence 1024 BP; 323 A; 215 C; 262 G; 224 T; 0 other;

```
Query Match      100.0%; Score 1024; DB 21; Length 1024;
```

Best Local Similarity	100.0%	Pred. No. 1.3e-263;
-----------------------	--------	---------------------

Matches 1024; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ACCGCTCGTCCGCGCGGCTTGAGCCGCGGGGAGCGCGCAATTGCTCGGCGCGCGG	60
Db	1	ACCGCTCGTCCGCGCGGCTTGAGCGCCCGGGGAGCGCGCAATTGCTCGGCGCGCGG	60
Qy	61	GGGGCGGCTCCCGGCGCATCTTCGCGCGCAACAAGGACTTACCAGGAAGGGAGCGGCTGGG	120
Db	61	GGGGCGGCTCCCGGCGCATCTTCGCGCGCAACCAGGACTTACCAGGAAGGGGAGCGGCTGGG	120
Qy	121	ATGGCGGCTCCGCGGCGCCCGCGAGTACAAAGCGGGCGACTTGGTCTTCGCCAAGATGAAG	180
Db	121	ATGGCGGCTCCGCGGCGCCCGCGAGTACAAAGCGGGCGACTTGGTCTTCGCCAAGATGAAG	180
Qy	181	GGCTACCCGCACTGGCGCGGCCGCGATTGATGAATCTCCAGAGGCGCTGTGAGAGCTTCCA	240
Db	181	GGCTACCCGCACTGGCGCGGCCGCGATTGATGAATCTCCAGAGGCGCTGTGAGAGCTTCCA	240
Qy	241	GCAACAAGTATCTATCTTCTTTTGGCACCCATGAATCTAGTATTCTAGGTCCCAA	300
Db	241	GCAACAAGTATCTATCTTCTTTTGGCACCCATGAATCTAGTATTCTAGGTCCCAA	300
Qy	301	GACCTTTTCCATATAAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACAACCGGAAGGA	360
Db	301	GACCTTTTCCATATAAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACAACCGGAAGGA	360
Qy	361	TTTAAACGAAGGATTGTGGGAAATAGAAATAACCCAGGAGTAAAGTTTACTTGGCTACCAG	420
Db	361	TTTAAACGAAGGATTGTGGGAAATAGAAATAACCCAGGAGTAAAGTTTACTTGGCTACCAG	420
Qy	421	GCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAGGTGGAAATATCTGCAGATGCA	480
Db	421	GCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAGGTGGAAATATCTGCAGATGCA	480
Qy	481	AGCAGTGAGGAAGAGGTGATAGCTTGAAGAGATGGAAGGCGAAAGCGAAAGAAATGAA	540
Db	481	AGCAGTGAGGAAGAGGTGATAGCTTGAAGAGATGGAAGGCGAAAGCGAAAGAAATGAA	540
Qy	541	AAAGCAGGCTCAAAACGGAAAAAGTCATATCTTCAAGAAATCCTCTTAAACAGTCCCGG	600
Db	541	AAAGCAGGCTCAAAACGGAAAAAGTCATATCTTCAAGAAATCCTCTTAAACAGTCCCGG	600
Qy	601	AAATCTCAGGAGATGAAGATGACAAAGACTGCAAGAGAGGAAAAACAAAGCAGCTCT	660
Db	601	AAATCTCAGGAGATGAAGATGACAAAGACTGCAAGAGAGGAAAAACAAAGCAGCTCT	660
Qy	661	GAGGTTGGAGATCGCGGCAACACACAAGAACCAACTTCAGACTTTCGAGAAACCCAGT	720
Db	661	GAGGTTGGAGATCGCGGCAACACACAAGAACCAACTTCAGACTTTCGAGAAACCCAGT	720
Qy	721	GAAAGGACCTAACTACCAATAATGAATGCTGCATATTAAGAGAAACCAAGAAAGTTATA	780
Db	721	GAAAGGACCTAACTACCAATAATGAATGCTGCATATTAAGAGAAACCAAGAAAGTTATA	780
Qy	781	TGTTTTGGTTGCTTAATATCTTGGATTGATATGAACCAACACATAGTCTCTTGTTCAT	840
Db	781	TGTTTTGGTTGCTTAATATCTTGGATTGATATGAACCAACACATAGTCTCTTGTTCAT	840
Qy	841	TGACAGAACCCGAGTTTGTATGATCAATTAATCATATTCCTCTCTGTTGTGTTTCGGGGG	900



CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.

XX  
SQ Sequence 1973 BP; 627 A; 367 C; 420 G; 559 T; 0 other;

Query Match 98.8%; Score 1012; DB 22; Length 1973;

Best Local Similarity 99.9%; Pred. No. 2.9e-260;

Matches 1023; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 2 CCGCTCGTCCCGCCGCTTGAAGCCCGCGGAGCGC-GCGCAATTGCTGGCCCGCGGG 60  
Db |||||  
Qy 36 CCGCTCGTCCCGCCGCTTGAAGCCCGCGGAGCGCGGCAATTGCTGGCCCGCGGG 95  
Db |||||  
Qy 61 GGGGGCGGCTCCCGGCACTTCGCGGCGACCAAGAGCTACCAAGAGGGGAGCGGCTGG 120  
Db |||||  
Qy 96 GGGGGCGGCTCCCGGCACTTCGCGGCGACCAAGAGCTACCAAGAGGGGAGCGGCTGG 155  
Db |||||  
Qy 121 ATGGCGGCTCCCGGCGCCCGCGAGTACAAAGCGCGGCACTTGTCTTCGCCCAAGATGAAG 180  
Db |||||  
Qy 156 ATGGCGGCTCCCGGCGCCCGCGAGTACAAAGCGCGGCACTTGTCTTCGCCCAAGATGAAG 215  
Db |||||  
Qy 181 GGCTACCGGCACTGGCGGCGCCGAGTATGAATCCCGAGAGCGGCTGTGAAGCCTCCA 240  
Db |||||  
Qy 216 GGCTACCGGCACTGGCGGCGCCGAGTATGAATCCCGAGAGCGGCTGTGAAGCCTCCA 275  
Db |||||  
Qy 241 GCAACAGATATCTATCTCTCTTTTGGCACCCATGAACTGCAATCTTAGTCTCCAAA 300  
Db |||||  
Qy 276 GCAACAGATATCTATCTCTCTTTTGGCACCCATGAACTGCAATCTTAGTCTCCAAA 335  
Db |||||  
Qy 301 GACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACAACGGAAGGA 360  
Db |||||  
Qy 336 GACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACAACGGAAGGA 395  
Db |||||  
Qy 361 TTTAACGAAGATTGTGGAAATAGAAATACCCAGAGTAAAGTTTACTGGCTACCAAG 420  
Db |||||  
Qy 396 TTTAACGAAGATTGTGGAAATAGAAATACCCAGAGTAAAGTTTACTGGCTACCAAG 455  
Db |||||  
Qy 421 GCAATTCAGCAACAGCTCTTCAGAACTGAGGAGAGAGTGAATATCTGCAGATGCA 480  
Db |||||  
Qy 456 GCAATTCAGCAACAGCTCTTCAGAACTGAGGAGAGAGTGAATATCTGCAGATGCA 515  
Db |||||  
Qy 481 AGCAGTGAAGAGAGGTGATAGAGTGAAGAGATGMAAAGCAAAAGAAAGAAATGAA 540  
Db |||||  
Qy 516 AGCAGTGAAGAGAGGTGATAGAGTGAAGAGATGMAAAGCAAAAGAAAGAAATGAA 575  
Db |||||  
Qy 541 AAAGCAGGCTCAAAACCGGAAAGTCTATATCTTCAAGAAATCTCTTAAACAGTCCCGG 600  
Db |||||  
Qy 576 AAAGCAGGCTCAAAACCGGAAAGTCTATATCTTCAAGAAATCTCTTAAACAGTCCCGG 635  
Db |||||  
Qy 601 AAATCTCAGAGATGAAGTGAACAAGACTGCAAGAGAGGAAAGCAAAAGCAGTCT 660  
Db |||||  
Qy 636 AAATCTCAGAGATGAAGTGAACAAGACTGCAAGAGAGGAAAGCAAAAGCAGTCT 695  
Db |||||  
Qy 661 GAGGCTGAGATGCGGCAACGACACAAAGAAACAACCTTCAGACTTGCAGAAACCAAGT 720  
Db |||||  
Qy 696 GAGGCTGAGATGCGGCAACGACACAAAGAAACAACCTTCAGACTTGCAGAAACCAAGT 755  
Db |||||  
Qy 721 GAAGGACCTTAACCTACCAATATGAATGCTGCATATTAAGAGAAACCAAGAAAGTTATA 780  
Db |||||  
Qy 756 GAAGGACCTTAACCTACCAATATGAATGCTGCATATTAAGAGAAACCAAGAAAGTTATA 815  
Db |||||  
Qy 781 TGTGTTGTTGTTAAATCTTGAATTTGATATGAACCAACATAGTCTCTTGTGTGCAT 840  
Db |||||  
Qy 816 TGTGTTGTTGTTAAATCTTGAATTTGATATGAACCAACATAGTCTCTTGTGTGCAT 875  
Db |||||  
Qy 841 TGACAGAACCCCGCTTGTATGTATACATTTATTCATATTCCTCTCTGTTGTTCGGGGGG 900  
Db |||||  
Qy 876 TGACAGAACCCCGCTTGTATGTATACATTTATTCATATTCCTCTCTGTTGTTCGGGGGG 935  
Db |||||  
Qy 901 AAAAGACATTTAGCCCTTTTAAAGAGTACTGATTTAAATTTCAATGTTATTTGGTTGCAT 960  
Db |||||

Db |||||  
Qy 936 AAAAGACATTTAGCCCTTTTAAAGAGTACTGATTTAAATTTTCATGTTATTTGGTTGCAT 995  
Db |||||  
Qy 961 GAGATTGGCCCTTAACCACTAAGGATTTATCAAGATTTTGGCGAGACTTATACATGTTCTAG 1020  
Db |||||  
Qy 996 GAAATTGGCCCTTAACCACTAAGGATTTATCAAGATTTTGGCGAGACTTATACATGTTCTAG 1055  
Db |||||  
Qy 1021 GATC 1024  
Db |||||  
Qy 1056 GATC 1059  
Db |||||

## RESULT 3

AAF76854

ID AAF76854 standard; cDNA; 1898 BP.

XX AAF76854;

XX AAF76854;

XX 14-MAY-2001 (first entry)

XX Human secreted protein cDNA #12.

XX Human; secreted protein; immunomodulatory; antisclerotic;

XX dermatological; immunosuppressive; antiinflammatory; anti-HIV;

XX immunostimulant; cytostatic; cardiant; vascular; anti-angiogenic;

XX ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;

XX antihelmithers; antiparkinsonian; antimicrobial; vulnery; gene therapy;

XX immune disorder; hyperproliferative; cardiovascular; angiogenic;

XX neurological; infection; ss.

XX Homo sapiens.

XX WO200112776-A2.

XX 22-FEB-2001.

XX 15-AUG-2000; 2000WO-US22350.

XX 16-AUG-1999; 99US-0148759.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Shi Y, Young PE, Ebner R, Soppet DR, Ruben SM;

XX WPI: 2001-244245/25.

XX P-PSDB; AAB70073.

XX Nucleic acids encoding 18 human secreted polypeptides, useful for

XX preventing, diagnosing and/or treating e.g. cancers, Parkinson's

XX disease and diabetic retinopathy -

XX Claim 1; Page 355-356; 380pp; English.

XX The present sequence is one of 18 nucleic acid molecules encoding novel

XX human secreted proteins. The nucleic acids and proteins may be used in

XX the prevention, diagnosis and treatment of diseases including immune

XX disorders (e.g. multiple sclerosis, systemic lupus erythematosus and

XX human immunodeficiency virus (HIV) infections), hyperproliferative

XX disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases

XX (e.g. Scimitar syndrome, Chaga's cardiomyopathy and coronary

XX arteriosclerosis), angiogenic disorders (e.g. corneal graft

XX neovascularisation and diabetic retinopathy), neurological disorders

XX (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),

XX infectious diseases and/or for promoting wound healing, regeneration

XX and/or chemotaxis. The nucleic acid molecules may also be used as DNA

XX probes in diagnostic assays to detect and quantitate the presence of

XX similar nucleic acid sequences in samples. The polypeptides may also be

XX used as antigens in the production of antibodies and in assays to

XX identify modulators of protein expression and activity.

XX SQ Sequence 1898 BP; 609 A; 363 C; 398 G; 520 T; 8 other;

Query Match

97.7%; Score 1000; DB 22; Length 1898;





```
QY 119 GGATGCGCGCTCCGCGCCCGCGAGTACAAAGCGCGGACCTGTCTTCGCAAGATGA 178
Db 122 GGATGCGCGCTCCGCGCCCGCGAGTACAAAGCGCGGACCTGTCTTCGCAAGATGA 181
QY 179 AGGCTACCCGCACTGGCGCGCCCGGATGATGAAGTCCAGAGGCGCTGTGAAGCTC 238
Db 182 AGGCTACCCGCACTGGCGCGCCCGGATGATGAAGTCCAGAGGCGCTGTGAAGCTC 241
QY 239 CAGCAAAACAAGTATCTATCTCTTTTGGGCAACCATGAAACTGCATTTCTAGGTCCA 298
Db 242 CAGCAAAACAAGTATCTATCTCTTTTGGGCAACCATGAAACTGCATTTCTAGGTCCA 301
QY 299 AAGACCTTTTCCATATAGAGTACAAAGACAGTTTGGAAAGTCAACAAACGGAAG 358
Db 302 AAGACCTTTTCCATATAGAGTACAAAGACAGTTTGGAAAGTCAACAAACGGAAG 361
QY 359 GATTTAACAAGGATTTGGGAAATAGAAAATAAACCAGGAGTAAGTTTACTGCTACC 418
Db 362 GATTTAACAAGGATTTGGGAAATAGAAAATAAACCAGGAGTAAGTTTACTGCTACC 421
QY 419 AGGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTGGAATATCTGCAGATG 478
Db 422 AGGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTGGAATATCTGCAGATG 481
QY 479 CAAGCAGTGAAGGAGTGTAGAGTAGAAGAGTGAAGAGGCAAGAAAGAAATG 538
Db 482 CAAGCAGTGAAGGAGTGTAGAGTAGAAGAGTGAAGAGGCAAGAAAGAAATG 541
QY 539 AAAAGCAGGCTCAAAACGGAAGTCTATATCTTCAAGAAATCTCTTAAACAGTCC 598
Db 542 AAAAGCAGGCTCAAAACGGAAGTCTATATCTTCAAGAAATCTCTTAAACAGTCC 601
QY 599 GGAATCTCCAGAGATGAAGATGACAAAGCTGCAAGAGAGAGGAAACAAACAGCT 658
Db 602 GGAATCTCCAGAGATGAAGATGACAAAGCTGCAAGAGAGAGGAAACAAACAGCT 661
QY 659 CTGAGGTGAGATGCGGCAACGACACAAAGAAACAACTTCAGACTTGCAGAAACCA 718
Db 662 CTGAGGTGAGATGCGGCAACGACACAAAGAAACAACTTCAGACTTGCAGAAACCA 721
QY 719 GTGAAGGACCTAATACATATGATGATGATGAACCAACATAGTCTCTGTTGC 838
Db 722 GTGAAGGACCTAATACATATGATGATGATGAACCAACATAGTCTCTGTTGC 841
QY 779 TATGTTTGTCTGTAATATCTTGGATTTGATGATGAACCAACATAGTCTCTGTTGC 888
Db 782 TATGTTTGTCTGTAATATCTTGGATTTGATGATGAACCAACATAGTCTCTGTTGC 891
QY 839 ATTGACAGAACCCAGTTTGTATGATGATGATGATGATGATGATGATGATGATGAT 898
Db 842 ATTGACAGAACCCAGTTTGTATGATGATGATGATGATGATGATGATGATGATGAT 901
QY 899 GGAAGAGACATTTAGCTTTTAAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTA 958
Db 902 GGAAGAGACATTTAGCTTTTAAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTA 961
QY 959 ATGAAGTTCCTTAAACCACTAAGATTTATCAAGATTTTGGCAGACTTATACATGCT 1018
Db 962 ATGAAGTTCCTTAAACCACTAAGATTTATCAAGATTTTGGCAGACTTATACATGCT 1021
QY 1019 AGGATC 1024
Db 1022 AGGATC 1027
```

## RESULT 6

ABS70439

ID ABS70439 standard; cDNA; 4487 BP.

XX

AC ABS70439;

XX

DT 27-NOV-2002 (first entry)

XX

DE Human bone remodelling gene #96.

XX Bone remodelling; osteoporosis; human; gene; ss.

XX Homo sapiens.

XX US6426186-B1.

XX 30-JUL-2002.

XX 18-JAN-2000; 2000US-0484970.

XX 18-JAN-2000; 2000US-0484970.

XX (INCY-) INCYTE GENOMICS INC.

XX Jones KA, Volkmut W, Walker MG;

XX WPI; 2002-673014/72.

XX A combination of polynucleotides which are co-expressed with genes

XX known to be involved in bone remodelling and osteoporosis are useful in

XX an array for the diagnosis of bone remodelling and osteoporosis

XX associated disorders

XX Claim 1; Column 263-268; 206pp; English.

XX The invention relates to a combination comprising a number of

XX substantially purified and isolated polynucleotides which are

XX co-expressed with genes known to be involved in bone remodelling and

XX osteoporosis. The invention is used to diagnose disorders associated

XX with bone remodelling or osteoporosis. ABS70344-ABS70512 represent

XX human bone remodelling genes of the invention.

XX Sequence 4487 BP; 1457 A; 826 C; 887 G; 1314 T; 3 other;

SQ

Query Match 94.6%; Score 969; DB 24; Length 4487;

Best Local Similarity 99.0%; Pred. No. 1.4e-248;

Matches 1017; Conservative 0; Mismatches 5; Indels 5; Gaps 4;

QY 1 ACCGCTCTCGCGCGCTTGAAGCCCGCGGAGCGCGCAATTCGTCGCGCGCGG 60

Db 4 ACCGCTCTCGCGCGCTTGAAGCCCGCGGAGCGCGCAATTCGTCGCGCGCGG 61

QY 61 GGGCGGCGC--TCCCGGCATCTTCGCGCGACCAAGGACTACCAGAAAGGAGCGGCTG 118

Db 62 GGGCGGCGCATCCCGGAGTCTTCGCGCGACCAAGGACTACCAGAAAGGAGCGGCTG 121

QY 119 GGATGCGCGCTCCGCGCCCGCGAGTACAAAGCGGCGGACCTGCTTCGCAAGATGA 178

Db 122 GGATGCGCGCTCCGCGCCCGCGAGTACAAAGCGGCGGACCTGCTTCGCAAGATGA 181

QY 179 AGGCTACCCGCACTGGCGCGCCCGGATGATGAATCCAGAGGCGCTGTGAAGCTC 238

Db 182 AGGCTACCCGCACTGGCGCGCCCGGATGATGAATCCAGAGGCGCTGTGAAGCTC 241

QY 239 CAGCAAAACAAGTATCTATCTCTTTTGGCACCACCATGAACTGCATTTCTAGGTCCA 298

Db 242 CAGCAAAACAAGTATCTATCTCTTTTGGCACCACCATGAACTGCATTTCTAGGTCCA 301

QY 299 AAGACCTTTTCCATATAGAGTACAAAGACAGTTTGGAAAGTCAACAAACGGAAG 358

Db 302 AAGACCTTTTCCATATAGAGTACAAAGACAGTTTGGAAAGTCAACAAACGGAAG 361

QY 359 GATTTAACAAGGATTTGGGAAATAGAAAATAAACCAGGAGTAAGTTTACTGCTACC 418

Db 362 GATTTAACAAGGATTTGGGAAATAGAAAATAAACCAGGAGTAAGTTTACTGCTACC 421

QY 419 AGGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTGGAATATCTGCAGATG 478

Db 422 AGGCAATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTGGAATATCTGCAGATG 481

QY 479 CAAGCAGTGAAGGAGTGTAGAGTAGAAGAGTGAAGAGGCAAGAAAGAAATG 538

```
XX PA (CORI-) CORIXA CORP.
XX PI Benson DR, Lodes MJ, Mitcham JL, King GE;
XX XX WPI; 2003-147101/14.
XX XX
XX XX Determining presence or absence of cancer in patient by contacting
XX PT patient sample with oligonucleotide that hybridises to polynucleotide
XX PT encoding ovarian carcinoma antigen, and detecting amount of
XX PT hybridisation in sample
XX XX
XX PS Example 1; Fig 16I; 152pp; English.
XX XX
XX CC The present invention relates to compositions and methods for
XX CC determining presence or absence of cancer in patient. The method
XX CC comprising contacting a biological sample with an oligonucleotide
XX CC that hybridises to partial polynucleotide sequence encoding for
XX CC human ovarian carcinoma antigen, or its complement. The method and
XX CC compositions are useful for the therapy and diagnosis of ovarian
XX CC cancer. The method can be used to identify tumour antigens that
XX CC are secreted from ovarian carcinoma and/or other tumours. Effective
XX CC cancer detection is achieved using the method of the invention.
XX CC ABX72720-ABX73050 represent partial cDNA sequences encoding human
XX CC ovarian carcinoma antigens.
XX XX
XX SQ Sequence 1664 BP; 566 A; 280 C; 319 G; 499 T; 0 other;

Query Match 80.9%; Score 828.4; DB 25; Length 1664;
Best Local Similarity 98.7%; Pred. No. 3.6e-211;
Matches 835; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 179 AGGCTACCCGCACTGGCCGCGGATTTGATGAATCCACAGAGGGCGCTGTGAAGCCTC 238
DB 17 AGGCGCGCCAGTCGGGGTAGCCCGGATTTGATGAATCCACAGAGGGCGCTGTGAAGCCTC 76
QY 239 CAGCAAAACAAGATCTCTATCTCTTTTGGCACCCATGAATCTGCAATTTCTAGGTCCCA 298
DB 77 CAGCAAAACAAGATCTCTATCTCTTTTGGCACCCATGAATCTGCAATTTCTAGGTCCCA 136
QY 299 AAGACCTTTTCCATATAAGAGTACAAAGCAAGTTTGGAAAGTCAAAACAGGAAAG 358
DB 137 AAGACCTTTTCCATATAAGAGTACAAAGCAAGTTTGGAAAGTCAAAACAGGAAAG 196
QY 359 GATTTAACGAAGAGTCTGGGAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACC 418
DB 197 GATTTAACGAAGAGTCTGGGAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACC 256
QY 419 AGCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAGGTGGAAATCTGCAAGTG 478
DB 257 AGCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAGGTGGAAATCTGCAAGTG 316
QY 479 CAAGCAGTGAAGAGAGGTGATAGATGAGAGAGTGAAGAGTGAAGAGGCAAAAGAAATG 538
DB 317 CAAGCAGTGAAGAGAGGTGATAGATGAGAGAGTGAAGAGTGAAGAGGCAAAAGAAATG 376
QY 539 AAAAAAGCAGGCTCAAAAAGGAAAGTCATATCTTCAAGAAATCTCTTAAACAGTCCC 598
DB 377 AAAAAAGCAGGCTCAAAAAGGAAAGTCATATCTTCAAGAAATCTCTTAAACAGTCCC 436
QY 599 GGAATCTCCAGGAGATGAAGATGACAAAGACTGCAAAAGAGGAGGAAACAAAGCAGCT 658
DB 437 GGAATCTCCAGGAGATGAAGATGACAAAGACTGCAAAAGAGGAGGAAACAAAGCAGCT 496
QY 659 CTGAGGTGAGATCGGGCAACGACACAAAGAACACAACTTCAGATTCGAGAAACCA 718
DB 497 CTGAGGTGAGATCGGGCAACGACACAAAGAACACAACTTCAGATTCGAGAAACCA 556
QY 719 GTGAGGGCCTTAACCTACCATAATGATGCTGATATTTAGAGNAACCAAGAGGTTA 778
DB 557 GTGAGGGCCTTAACCTACCATAATGATGCTGATATTTAGAGNAACCAAGAGGTTA 616
QY 779 TATGTTGGTGTCTAATATTTCTTGGATTTGATATGAACCAACACATAGTCTTGTGTC 838
```

```
DB 617 TATGTTGGTGTCTAATATTTCTTGGATTTGATGAACCAACACATAGTCTTGTGTC 676
QY 839 ATTGACAGAACCCAGTTTGTATGATCATATTTCTCTCTCTGTTGTTGGGG 898
DB 677 ATTGACAGAACCCAGTTTGTATGATCATATTTCTCTCTCTGTTGTTGGGG 736
QY 899 GGAAGAGACATTTTAGCCTTTTAAAGTTACTGATTTAAATTTCAATTTGTTGTC 958
DB 737 GGAAGAGACATTTTAGCCTTTTAAAGTTACTGATTTAAATTTCAATTTGTTGTC 796
QY 959 ATGAAGTTGCCCTTAACCACTAAGGATTAAGAGTTTATCAAGATTTTTCGAGACTTATACATGCT 1018
DB 797 ATGAAGTTGCCCTTAACCACTAAGGATTAAGAGTTTATCAAGATTTTTCGAGACTTATACATGCT 856
QY 1019 AGGATC 1024
DB 857 AGGATC 862

RESULT 10
ABX72983
ID ABX72983 standard; cDNA; 1664 BP.
XX AC ABX72983;
XX XX
XX DT 14-MAR-2003 (first entry)
XX DE Human ovarian carcinoma antigen partial cDNA sequence #249.
XX KW Human; cancer detection; ovarian carcinoma antigen; ovarian cancer;
XX KW tumour antigen; tumour; OV2; OV3; OV6; OV9; OV10; OV12; OV14; OV17;
XX KW OV18; OV23; OV24; OV27; OV41; OV54; OV57; gene; ss.
XX OS Homo sapiens.
XX XX
XX PN US6468758-B1.
XX PD 22-OCT-2002.
XX XX
XX PF 16-SEP-1999; 99US-0397787.
XX PR 23-SEP-1998; 98US-0159320.
XX PR 08-FEB-1999; 99US-0246429.
XX XX
XX PA (CORI-) CORIXA CORP.
XX XX
XX PI Benson DR, Lodes MJ, Mitcham JL, King GE;
XX XX WPI; 2003-147101/14.
XX XX
XX PT Determining presence or absence of cancer in patient by contacting
XX PT patient sample with oligonucleotide that hybridises to polynucleotide
XX PT encoding ovarian carcinoma antigen, and detecting amount of
XX PT hybridisation in sample
XX XX
XX PS Example 1; Column 193-196; 152pp; English.
XX XX
XX CC The present invention relates to compositions and methods for
XX CC determining presence or absence of cancer in patient. The method
XX CC comprising contacting a biological sample with an oligonucleotide
XX CC that hybridises to partial polynucleotide sequence encoding for
XX CC human ovarian carcinoma antigen, or its complement. The method and
XX CC compositions are useful for the therapy and diagnosis of ovarian
XX CC cancer. The method can be used to identify tumour antigens that
XX CC are secreted from ovarian carcinoma and/or other tumours. Effective
XX CC cancer detection is achieved using the method of the invention.
XX CC ABX72720-ABX73050 represent partial cDNA sequences encoding human
XX CC ovarian carcinoma antigens.
XX XX
XX SQ Sequence 1664 BP; 566 A; 280 C; 319 G; 499 T; 0 other;

Query Match 80.9%; Score 828.4; DB 25; Length 1664;
```

	Best Local Similarity	98.7%; Pred. No. 3.6e-211;	Matches 835; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY	179	AGGGCTACCCGACTCGCCGCCCGATTCATGAATCCCGACAGGGCGCTGTGAAGCCTC	238
Db	17	AGGGCCGCAGTCCGGGTAGCCCGNTGATGAACTCCACAGAGGGCGCTGTGAAGCCTC	76
QY	239	CAGCAACAAGTATCTCATCTCTTTTTTTGGCAACCATGAAACTGCAATTCCTAGTGCCCA	298
Db	77	CAGCAACAAGTATCTCATCTCTTTTTTTGGCAACCATGAAACTGCAATTCCTAGTGCCCA	136
QY	299	AAGACCTTTTTCCATAAAGGCTACAAAGACAAGTTTGGAAAGTCAACAAACGGAAG	358
Db	137	AAGACCTTTTTTCCATAAAGGAGTACAAAGACAAGTTTGGAAAGTCAACAAACGGAAG	196
QY	359	GATTTTAACGAAAGGATTTGTGGGAATAGAAAAATAACCCAGGAGTAAAGTTTACTGGCTACC	418
Db	197	GATTTTAACGAAAGGATTTGTGGGAATAGAAAAATAACCCAGGAGTAAAGTTTACTGGCTACC	256
QY	419	AGGCNATTCAGCAACACAGAGCTCTTCAGAAACTGAGGGAGAAGGTGGAANAATCTGCAGATG	478
Db	257	AGGCNAATTCAGCAACACAGAGCTCTTCAGAAACTGAGGGAGAAGGTGGAANAATCTGCAGATG	316
QY	479	CAAGCAGTGAAGGAAGAAGGTGATAGAGTGAAGAAGATGAAAGGCAAAAGAAAGAATG	538
Db	317	CAAGCAGTGAAGGAAGAAGGTGATAGAGTGAAGAAGATGGAAGAAGGCAAAAGAAAGAATG	376
QY	539	AAAAAGCAGGCTCAAAAACGGAAAAAGTCATPATCTTCAAAGAAATCTCTTAAACAGTCCC	598
Db	377	AAAAAGCAGGCTCAAAAACGGAAAAAGTCATPATCTTCAAAGAAATCTCTTAAACAGTCCC	436
QY	599	GGAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAGAAGAGGAAACAAAGACGCT	658
Db	437	GGAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAGAAGAGGAAACAAAGACGCT	496
QY	659	CTGAGGGTGGAGATGCGGGCCAAACGACAAAGAAACAACTTCAGACTTCGAGAAAACCA	718
Db	497	CTGAGGGTGGAGATGCGGGCCAAACGACAAAGAAACAACTTCAGACTTCGAGAAAACCA	556
QY	719	GTGAAGGGAACCTACTACCAATATGAATGCTGCATATTAAGAGAAACCAACGAAGAGTTA	778
Db	557	GTGAAGGGAACCTACTACCAATATGAATGCTGCATATTAAGAGAAACCAACGAAGAGTTA	616
QY	779	TATGTTTGGTGTCTATAATCTTGGATTTGATATCAACCAACATAGTCCCTGTTGTC	838
Db	617	TATGTTTGGTGTCTATAATCTTGGATTTGATATCAACCAACATAGTCCCTGTTGTC	676
QY	839	ATTGACAAACCCAGCTTTGTATGTACATTAATTCATATTCCTCTCTGTTGTTTTCGGGG	898
Db	677	ATTGACAAACCCAGCTTTGTATGTACATTAATTCATATTCCTCTCTGTTGTTTTCGGGG	736
QY	899	GGAAAGACATTTTAGGCTTTTTTAAAGTACTGATTTTAATTTCAATGTTATTTGGTTGC	958
Db	737	GGAAAGACATTTTAGGCTTTTTTAAAGTACTGATTTTAATTTCAATGTTATTTGGTTGC	796
QY	959	ATGAAGTTGCCCTTTAAACCACTAAGGATTATCAAGATTTTTTGGCAGACTTATACATGTCT	1018
Db	797	ATGAAGTTGCCCTTTAAACCACTAAGGATTATCAAGATTTTTTGGCAGACTTATACATGTCT	856
QY	1019	AGGATC 1024	
Db	857	AGGATC 862	

RESULT 11

ABL35012

ID ABL35012 standard; cDNA; 2920 BP.

XX 25

AC ABL39

XX  
XX  
XX

DT 04-A)

U4-A1  
DI XX

XX	Human; rat; mouse; skin cell; skin wound; cancer; growth defect; developmental defect; inflammatory disease; dermatological; vulnary; immunomodulator; anti-inflammatory; cytostatic; neuroprotective; gene; ss.
XX	Rattus sp.
XX	WO200190357-A1.
XX	29-NOV-2001.
XX	24-MAY-2001; 2001WO-N200099.
XX	24-MAY-2000; 2000US-206650P.
XX	25-JUL-2000; 2000US-221232P.
XX	(GENE-) GENESIS RES & DEV CORP LTD.
XX	Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG, Kumble KD; WPI: 2002-122020/16.
XX	P-PSDB; ABB72327.
XX	New polynucleotides and polypeptides encoded by the polynucleotides isolated from skin cells, useful for treating skin wounds, cancers, growth and developmental defects, inflammatory diseases, or for modulating immune responses -
XX	Claim 1; Page 334-335; 466pp; English.
XX	The present invention provides the protein and coding sequences of cDNAs isolated from human, murine and rat skin cell libraries. The sequences can be used in the development of therapeutic agents useful in the treatment of skin diseases, including skin wounds, cancer, growth defects, developmental defects and inflammatory diseases. The proteins have important roles in the induction of hair growth, cell proliferation and cell-cell interaction, in maintaining tissue integrity, in wound healing and in modulating immune responses. The present sequence is a cDNA of the invention.
XX	Sequence 2920 BP; 877 A; 600 C; 652 G; 791 T; 0 other;
XX	Query Match 76.0%; Score 778; DB 24; Length 2920;
XX	Best Local Similarity 89.1%; Pred. No. 1.4e-197;
XX	Matches 912; Conservative 0; Mismatches 95; Indels 17; Gaps 6;
Qy	2 CCGCTCTCGCCCGCGGTTGAGGCCCGCGGGAGCGC-GGCGAATTCGTGCGCCCGCGGG 60
Db	164 CCGCTCAGCGCCCGGCTCGAGGCCCGCGGGAAGCGCGCGTCTGCTCTCCCGCGGA 223
Qy	61 GGGCGCGCTCCCGGCATCTTCGGCGGACCAAGGACTACCGAGGGGAGCGGCTGGG 120
Db	224 GGGCGCGCTTCGGGCATCTTCGGCGGACCAAGGACCGCGGCGGGGAACCAACGGG 283
Qy	121 ATGCGCGCTCCGGCGCCCGCGAGTCAAAAGCGGCGACCTGGTCTTCGCCAAGATGAAG 180
Db	284 ATGGCGGCTCCGGCGCCCGCGAGTCAAAAGCGGAGACCTGGTCTTCGCCAATGAAG 343
Qy	181 GGCTACCCGCACTGGCGCGGCCGGATGTGAACCTCCAGAGGCGCTGTGAAGCCTTCCA 240
Db	344 GGCTACCCGCACTGGCGCGGCCGGATGTGAACCTCCAGAGGCGCGCTGAAGCCTTCCA 403
Qy	241 GCAACCAAGTATCCTATCTCTTTTGGCACCCATGAATCTGATTTCTAGGTCCTCCAAA 300
Db	404 GCAACCAAGTATCCTATCTCTTTTGGCACCCATGAATCTGATTTCTAGGTCCTTAAA 463
Qy	301 GACCTTTTCCATATAGGAGTCAAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAGGA 360
Db	464 GACCTTTTCCATATAGGAGTCAAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAGGA 523
Qy	361 TTTAACGAAGGATGTGGGAAATAGAAAAATAAACCCAGGAGTAAAGTTTATCTGGCTACCA 420

Db 524 TTATATCAAGGATTATGGGAAATTTGAAATAATCCAGGAGTGAAATTTACTGGGTACCAG 583  
Qy 421 GCAATTCAGCAACAGAGCTCTTCAGAAATCTGAGGAGAGGTGGAATACTGCAGATGCA 480  
Db 584 ACAATTCAGCAACAGAGCTCTTCAGAAATCTGAGGAGAGGTGGAATACTGCAGATGCA 643  
Qy 481 AGCAGTCAGGAAGAGTGTAGAGTAGAGAGATGGAAGGCAAAAGAAAGATGAA 540  
Db 644 AGCAGTCAGGAAGAGTGTAGAGTAGAGATGGAAGGCAAAAGAAAGATGAA 700  
Qy 541 AAAGCAGGCTCAAAACCGGAAAGTGTATATCTTCAAGAAATCTCTTAACAGTCCCG 600  
Db 701 AAAGCAGGCTCAAAACCGGAAAGTGTATATCTTCAAGAAATCTCTTAACAGTCCCG 760  
Qy 601 AAATCTCCAGGAGATGAAGATGACAAAGACTGCAAGAGAGGAGGAAACAAAGCAGTCT 660  
Db 761 AAATCTCCAGGAGATGAAGATGACAAAGACTGCAAGAGAGGAGGAAACAAAGCAGTCT 820  
Qy 661 GAGGTGGAGATGGGGCAACGACACAAAGAAACAACTTCAGACTTGCAGAAACAGT 720  
Db 821 GAGGTGGAGATGGGGCAACGACACAAAGAAACAACTTCAGACTTGCAGAAACAGT 880  
Qy 721 GAAGGACCTTAATACCTATGATGATGCTGCATATTAAGAAACCAACCAAGAGTTATA 780  
Db 881 GAAGGACCTTAATACCTATGATGATGCTGCATATTAAGAAACCAACCAAGAGTTATA 940  
Qy 781 TGTTGGTCTCTAAATATCTTGATTTGATGATGATGATGATGATGATGATGATGAT 840  
Db 941 TGTTGGTCTCTAAATATCTTGATTTGATGATGATGATGATGATGATGATGATGAT 900  
Qy 841 TGACAGAACCCAGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
Db 995 TGACAGAACCCAGTTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1050  
Qy 901 AAAGACATTTAGCTTTTAAAGTTACTGATTTAAATTTCAATTTTATGTTTGTGAT 960  
Db 1051 AAAGACATTTAGCTTTTAAAGTTACTGATTTAAATTTCAATTTTATGTTTGTGAT 1110  
Qy 961 GAAGTTGCCCTTAACCACTAAGGATTTCAAGATTTTGGCAGACTTATACATGCTAG 1020  
Db 1111 GAAGTTGCCCTTAACCACTAAGGATTTCAAGATTTTGGCAGACTTATACATGCTAG 1167  
Qy 1021 GATC 1024  
Db 1168 GCTC 1171

RESULT 12  
AAH05630  
ID AAH05630 standard; cDNA; 737 BP.  
XX  
AC AAH05630;  
XX  
DT 26-JUN-2001 (first entry)  
XX  
DE Human cDNA clone (5'-primer) SEQ ID NO:2465.  
XX  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX Homo sapiens.  
OS  
PN EP1074617-A2.  
XX  
PD 07-FEB-2001.  
XX  
PF 28-JUL-2000; 2000EP-0116126.  
XX  
PR 29-JUL-1999; 99JP-0248036.  
PR 27-AUG-1999; 99JP-0300253.  
PR 11-JAN-2000; 2000JP-0118776.  
PR 02-MAY-2000; 2000JP-0183767.  
PR 09-JUN-2000; 2000JP-0241899.  
XX

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;  
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602  
full-length cDNAs defined in the specification, and for the detection  
and/or diagnosis of the abnormality of the proteins encoded by the  
full-length cDNAs -

Claim 1; SEQ ID 2465; 2537pp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602  
full-length cDNAs defined in the specification. Where a primer set  
comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
to the complementary strand of a polynucleotide which comprises one of  
the 5602 nucleotide sequences defined in the specification, where the  
oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
of an oligonucleotide comprising a sequence complementary to the  
complementary strand of a polynucleotide which comprises a 5'-end  
sequence and an oligonucleotide comprising a sequence complementary to a  
polynucleotide which comprises a 3'-end sequence, where the  
oligonucleotide comprises at least 15 nucleotides and the combination of  
the 5'-end sequence/3'-end sequence is selected from those defined in  
the specification. The primer sets can be used in antisense therapy and  
in gene therapy. The primers are useful for synthesizing polynucleotides,  
particularly full-length cDNAs. The primers are also useful for the  
detection and/or diagnosis of the abnormality of the proteins encoded by  
the full-length cDNAs. The primers allow obtaining of the full-length  
cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
AAH98893 represent human amino acid sequences; and AAH13629 to AAH13632  
represent oligonucleotides, all of which are used in the exemplification  
of the present invention.

Sequence 737 BP; 227 A; 180 C; 204 G; 117 T; 9 other;

Query Match 60.9%; Score 623.6; DB 22; Length 737;  
Best Local Similarity 96.2%; Pred. No. 1.5e-156;  
Matches 676; Conservative 0; Mismatches 23; Indels 4; Gaps 4;

Qy 2 CGCTCGTCGCGCCGCGCTTTGAGCGCGCGGAGCGC-GCGCAATTCGTGCGGCCCGCGG 60  
Db 36 CGCTCGTCGCGCCGCGCTTTGAGCGCGCGGAGCGC-GCGCAATTCGTGCGGCCCGCGG 95  
Qy 61 GGGCGGCGCTCCCGGCATCTTCGCGCGACCAAGGACTACCAAGGAGGCGGCTGGG 120  
Db 96 GGGCGGCGCTCCCGGCATCTTCGCGCGACCAAGGACTACCAAGGAGGCGGCTGGG 155  
Qy 121 ATGGCGCGTCGCGGCCCGCGAGTACAAAGCGGGGACCTGGTCTTCGCCAAGATGAAG 180  
Db 156 ATGGCGCGTCGCGGCCCGCGAGTACAAAGCGGGGACCTGGTCTTCGCCAAGATGAAG 215  
Qy 181 GGTACCCGCTGCGCGGCCCGCGGATTTGATGAATCCAGAGGCGGCTGTGAGGCTCCA 240  
Db 216 GGTACCCGCTGCGCGGCCCGCGGATTTGATGAATCCAGAGGCGGCTGTGAGGCTCCA 275  
Qy 241 GCAAAACAAGTATCTATCTTTTGGCACCCATGAAATGCAATTTCTAGGTCCCAA 300  
Db 276 GCAAAACAAGTATCTATCTTTTGGCACCCATGAAATGCAATTTCTAGGTCCCAA 335  
Qy 301 GACCTTTTCCATATAAGAGTACAAAGCAAGATTGGAAAGTCAAAACAACGGAAGGA 360  
Db 336 GACCTTTTCCATATAAGAGTACAAAGCAAGATTGGAAAGTCAAAACAACGGAAGGA 395  
Qy 361 TTTAACGAAGGATTTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACCAG 420  
Db 396 TTTAACGAAGGATTTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACCAG 455  
Qy 421 GCAATTCAGCAACAGAGCTCTTTCAGAAACTGAGGGAGAAAGTGGAAATATCTGCAGATGCA 480

Db 456 GCAATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAAAGTGGAAATACTGCGAGATGCA 515  
Qy 481 AGCAGTGGAGAAAGGTGATAGACTAGAGAGATGGAAGAGGCAAAAGAAAGATGAA 540  
Db 516 AGCAGTGGAGAAAGGTGATAGACTAGAGAGATGGAAGAGGCAAAAGAAAGATGAA 575  
Qy 541 AAAGCAGGCTCAAAACGGAAGATCATATATCTTCAAAAGAAATCTCTAAACAGTCCCGG 600  
Db 576 AAAGCAGGCTCAAAACGGAAGAA-TCATATACTTCAAGAAATCTCTAAACAGTCCCGG 634  
Qy 601 AAATCTCCAGG-AGATGAAGATGACAAAGACTGCAAGAGAGGAAAAACAAAGCAGTCTC 659  
Db 635 AAATCTCCAGGAGATGAAGATGACAAAGACTGCCNAGAAANAAGAAAAACNAAAGCAGTCTC 694  
Qy 660 TGAGGCTGGAGATG-CGGGCAACGACACAAAGAAACACACTTC 701  
Db 695 TGAAGGTGGANATGCCGGCAACGACACNAGAAAAACACACTTC 737

## RESULT 13

AAAG5004

ID AAAG5004 standard; cDNA; 609 BP.

AC AAAG5004;

XX 13-NOV-2000 (first entry)

XX Human hepatoma-derived growth factor homologous polypeptide cDNA.

DE Human; hepatoma-derived growth factor homologous polypeptide;

XX huHDGFh; cell proliferation; haematopoiesis; angiogenesis;

KW lymphocyte proliferation; infection; autoimmune disease;

KW vascular disease; cancer; ss.

XX Homo sapiens.

XX Key

FH Location/Qualifiers

CDS 1..609

FT /\*tag= a

FT /product= "huHDGFh"

FT WO200037492-A2.

XX 29-JUN-2000.

XX 22-DEC-1999; 99WO-US30932.

XX 22-DEC-1998; 98US-0113344.

XX (ELIL ) LILLY &amp; CO ELI.

XX Na S;

XX PI

XX WPI; 2000-442638/38.

XX P-PSDB; AAB13521.

XX Novel human homologues of hepatoma-derived growth factor homologous  
PT nucleic acids, polypeptides useful as probes or amplification primers  
PT in the detection, quantitation or isolation of gene sequences or  
PT transcripts

XX Claim 3; Page 73-74; 75pp; English.

XX The present sequence is the human homologue of the hepatoma-derived  
CC growth factor coding sequence (huHDGFh). The protein produced from this  
CC gene is involved in tumour formation in some cells, and it is thought  
CC that the gene and protein will be useful in the diagnosis and treatment  
CC of infections, autoimmune disorders, vascular diseases and cancers. In  
CC addition, the gene can be used to produce transgenic animals which may be  
CC used as animal models for these diseases, and the protein can be used to  
CC induce cell proliferation, haematopoiesis, lymphocyte proliferation and  
CC angiogenesis.

SQ Sequence 609 BP; 223 A; 125 C; 159 G; 102 T; 0 other;

Query Match 59.5%; Score 609; DB 21; Length 609;

Best Local Similarity 100.0%; Pred. No. 1.1e-152; Length 609;

Matches 609; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 ATGGCGGTCCGCGGCCCGCGAGTACAAAGCGGGCGACCTTGCTTCGCGCAAGATGAAG 180

Db 1 ATGGCGGTCCGCGGCCCGCGAGTACAAAGCGGGCGACCTTGCTTCGCGCAAGATGAAG 60

Qy 181 GGCTACCCGCACTGGCGGCCCGCGAGTATGATGAACCTCCAGAGGGCGCTGTGAAGCTTCA 240

Db 61 GGCTACCCGCACTGGCGGCCCGCGAGTATGATGAACCTCCAGAGGGCGCTGTGAAGCTTCA 120

Qy 241 GCAACAAAGTATCTATCTCTCTTTTGGGCAACCATGAACTGCAATTTCTAGGTCCCAAA 300

Db 121 GCAACAAAGTATCTATCTCTCTTTTGGGCAACCATGAACTGCAATTTCTAGGTCCCAAA 180

Qy 301 GACCTTTTCCATATATAAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACGGAAGGA 360

Db 181 GACCTTTTCCATATATAAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACGGAAGGA 240

Qy 361 TTTAACGAAGATTGTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACCAG 420

Db 241 TTTAACGAAGATTGTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACCAG 300

Qy 421 GCAATTCCAGCAACAGAGCTCTTTCAGAACTCAGGAGAGAGTGGAAATCTGCAGATGCA 480

Db 301 GCAATTCCAGCAACAGAGCTCTTTCAGAACTCAGGAGAGAGTGGAAATCTGCAGATGCA 360

Qy 481 AGCAGTCCAGGAGAGGAGTGTAGAGTAGAAGAGATGGAAGAGGCAAAAGAAAGATGAA 540

Db 361 AGCAGTCCAGGAGAGGAGTGTAGAGTAGAAGAGATGGAAGAGGCAAAAGAAAGATGAA 420

Qy 541 AAAGCAGGCTCAAAACGGAAGAAAGTCTATATCTTCAAAAGAAATCTCTTAAACAGTCCCGG 600

Db 421 AAAGCAGGCTCAAAACGGAAGAAAGTCTATATCTTCAAAAGAAATCTCTTAAACAGTCCCGG 480

Qy 601 AAATCTCCAGGAGATGAGATGACAAAGACTGCAAGAGAGAGGAAACAAAGCAGCTCT 660

Db 481 AAATCTCCAGGAGATGAGATGACAAAGACTGCAAGAGAGAGGAAACAAAGCAGCTCT 540

Qy 661 GAGGTGGAGATGCGGCGCAACGACACACAGAAACACACTTTCAGACTTGCAGAAAAACAGT 720

Db 541 GAGGTGGAGATGCGGCGCAACGACACACAGAAACACACTTTCAGACTTGCAGAAAAACAGT 600

Qy 721 GAAGGGACC 729

Db 601 GAAGGGACC 609

## RESULT 14

AAAG79520

ID AAAG79520 standard; cDNA; 870 BP.

XX AC

XX AAAG79520;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #15324.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX PD 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2004, 08:55:40 ; Search time 78 Seconds  
(without alignments)  
5794.568 Million cell updates/sec

Title: US-09-787-328B-3

Perfect score: 1024

Sequence: 1 accgctcgccgccggctt.....actatacatgtctagatc 1024

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	
1	969	94.6	4487	4	US-09-484-970B-96	
2	828.4	80.9	1664	4	US-09-397-787-68	
3	828.4	80.9	1664	4	US-09-397-787-267	
C	4	471	46.0	645	3	US-09-328-111-318
5	193.4	18.9	2376	2	US-08-760-745-4	
6	188.6	18.4	869	2	US-08-760-745-2	
C	7	98.2	9.6	7218	1	US-08-232-463-14
8	50	4.9	696	3	US-09-461-697-193	
9	50	4.9	699	3	US-09-461-697-191	
10	50	4.9	717	3	US-09-461-697-189	
11	50	4.9	774	3	US-09-461-697-187	
12	50	4.9	819	3	US-09-461-697-185	
13	50	4.9	1280	4	US-09-220-132-155	
14	50	4.9	1669	3	US-09-461-697-184	
15	47.6	4.6	2223	1	US-08-257-073-4	
C	16	47.2	4.6	2791	4	US-09-570-367C-1
17	46.8	4.6	1298	4	US-08-971-089-1	
18	45.4	4.4	1194	3	US-08-929-329-3	
19	45.2	4.4	340	1	US-08-182-175A-104	
20	45.2	4.4	340	5	PCT-US92-06412-104	
C	21	45.2	4.4	929	4	US-09-671-317-14
22	45.2	4.4	1001	4	US-09-671-317-439	
C	23	44	4.3	44453	4	US-09-146-053-5
24	43.8	4.3	289	3	US-09-007-005-17	
25	43.8	4.3	289	3	US-09-244-796-17	
26	42.6	4.2	1298	3	US-08-948-705-3	
27	42.6	4.2	1298	4	US-09-510-543-3	

28	42.6	4.2	6755	3	US-08-931-999-4	Sequence 4, Appl1
29	42.4	4.1	282	3	US-09-461-697-205	Sequence 205, App
30	42.4	4.1	306	3	US-09-461-697-203	Sequence 203, App
31	42	4.1	43795	3	US-08-742-185-101	Sequence 101, App
32	41.8	4.1	2447	2	US-09-014-969-14	Sequence 14, Appl
33	41.6	4.1	376	2	US-08-623-906A-18	Sequence 18, Appl
34	41.6	4.1	4226	4	US-09-620-312D-480	Sequence 480, App
35	41.6	4.1	4766	5	PCT-US93-07261-10	Sequence 10, Appl
36	41.6	4.1	6002	4	US-09-345-882-4	Sequence 4, Appl1
37	41.4	4.0	189	3	US-09-461-697-213	Sequence 213, App
38	41.4	4.0	195	3	US-09-461-697-211	Sequence 211, App
39	41.4	4.0	213	3	US-09-461-697-209	Sequence 209, App
40	41.4	4.0	231	3	US-09-461-697-207	Sequence 207, App
41	41.4	4.0	152331	3	US-09-128-155-16	Sequence 16, App
42	41.4	4.0	1664976	4	US-08-916-421B-1	Sequence 1, Appl1
43	40.6	4.0	980	4	US-09-171-209-8	Sequence 8, Appl1
44	40.6	4.0	5394	3	US-08-688-376-1	Sequence 1, Appl1
45	40.4	3.9	1447	4	US-09-443-041A-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1

US-09-484-970B-96  
; Sequence 96, Application US/09484970B  
; Patent No. 6426186  
; GENERAL INFORMATION:  
; APPLICANT: Jones, Karen A.  
; APPLICANT: Volkmut, Wayne  
; TITLE OF INVENTION: BONE REMODELING GENES  
; FILE REFERENCE: PB-0014 US  
; CURRENT APPLICATION NUMBER: US/09/484,970B  
; CURRENT FILING DATE: 2000-01-18  
; NUMBER OF SEQ ID NOS: 172  
; SOFTWARE: PERL Program  
; SEQ ID NO 96  
; LENGTH: 4487  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6426186 330852.1CB1  
; NAME/KEY: unsure  
; LOCATION: 1985, 1992, 1997  
; OTHER INFORMATION: a, t, c, g, or other  
; US-09-484-970B-96

Query Match	94.6%	Score 969;	DB 4;	Length 4487;
Best Local Similarity	99.0%	Pred. No. 1.7e-270;		
Matches 1017;	Conservative 0;	Mismatches 5;	Indels 5;	Gaps 4;
Qy	1	ACCGCTCGTCGCGCCCGGCTTGAGCGCCGCGGGGAGCGCGCAATTTCGTGCGCCCGCGGG	60	
Db	4	ACCGCTCGTCGCGCCCGGCTTGAGCGCCGCGGGGAG-GCGCGCAATTTCGTGCGCCCGCGGG	61	
Qy	61	GGGGCGGCCC--TCGCGGCATCTTCGCGGGGACCAAGACTACCAAGAGGGGAGCGGCTG	118	
Db	62	GGGGCGGCCCATCCCGGCAGTCTTCGCGGGGACCAAGACTACCAAGAGGGGAGCGGCTG	121	
Qy	119	GGATGGCGGCTCGCGGCGCCCGGAGTACAAAGCGGCGACCTGCTTCGCCCAAGATGA	178	
Db	122	GGATGGCGGCTCGCGGCGCCCGGAGTACAAAGCGGCGACCTGCTTCGCCCAAGATGA	181	
Qy	179	AGGGCTACCGGCACTGGCGCGCCCGGATGATGAACCTCCAGAGGGCGCTGTGAAGCCTC	238	
Db	182	AGGGCTACCGGCACTGGCGCGCCCGGATGATGAACCTCCAGAGGGCGCTGTGAAGCCTC	241	
Qy	239	CAGCAAAAGATATCCTATCTTTTGGGACCCATGAACATGATTTCTAGGTCCTCA	298	
Db	242	CAGCAAAAGATATCCTATCTTTTGGGACCCATGAACATGATTTCTAGGTCCTCA	301	

Qy	299	AAGACCTTTTCCATATAAGAGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAG	358
Db	302	AAGACCTTTTCCATATAAGAGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAG	361
Qy	359	GATTTTAAAGGAAGTTGTGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTTCGGCTACC	418
Db	362	GATTTTAAAGGAAGTTGTGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTTCGGCTACC	421
Qy	419	AGGCAATTTCAGCAACAGAGCTCTTTTCAGAAATCTGAGGAGAGAGTGGAAATATCTCAGATG	478
Db	422	AGGCAATTTCAGCAACAGAGCTCTTTTCAGAAATCTGAGGAGAGAGTGGAAATATCTCAGATG	481
Qy	479	CAAGCAGTGGAGAAAGGTCATAGAGTAGAAGAAAGATGGAAAGGCCAAAGAAAGAAATG	538
Db	482	CAAGCAGTGGAGAAAGGTCATAGAGTAGAAGAAAGATGGAAAGGCCAAAGAAAGAAATG	541
Qy	539	AAAAAGCAGGCTCAAAACGGAAAAAGTCATATCTTCAAGAAATCTCTTAAACAGTCCC	598
Db	542	AAAAAGCAGGCTCAAAACGGAAAAAGTCATATCTTCAAGAAATCTCTTAAACAGTCCC	601
Qy	599	GGAAATCTCCAGGAGATGAAGTACAAAGACTGCAAGAGAGAGGAGAAAAACAAAGACAGCT	658
Db	602	GGAAATCTCCAGGAGATGAAGTACAAAGACTGCAAGAGAGAGGAGAAAAACAAAGACAGCT	661
Qy	659	CTGAGGGTGGAGATCGCGGCACAGACACAAAGACATTCAGACTTTCAGAGAAAAACCA	718
Db	662	CTGAGGGTGGAGATCGCGGCACAGACACAAAGACATTCAGACTTTCAGAGAAAAACCA	721
Qy	719	GTGAAGGGACCTTAACCTACCATAATGAATGCTGCATATTAAGAGAAAAACCAAA-GAAGGTT	777
Db	722	GTGAAGGGACCTTAACCTACCATAATGAATGCTGCATATTAAGAGAAAAACCAAGGAAGGTT	781
Qy	778	ATATGTTTGGTTGTCTAATATTTCTTGGATTTGATGAACCAACACATAGTCCCTTGTGTG	837
Db	782	ATATGTTTGGTTGTCTAATATTTCTTGGATTTGATGAACCAACACATAGTCCCTTGTGTG	841
Qy	838	CATTGACAGAACCCAGTTGTATGATCATATTAATTCATATTCCTCTCTGTGTGTGTTCCGG	897
Db	842	CATTGACAGAACCCAGTTGTATGATCATATTAATTCATATTCCTCTCTGTGTGTGTTCCGG	901
Qy	898	GGGAAAAGACATTTTAGCCTTTTTTAAAGTTACTGATTTAAATTCATGTATTATGTTG	957
Db	902	GGGAAAAGACATTTTAGCCTTTTTTAAAGTTACTGATTTAAATTCATGTATTATGTTG	961
Qy	958	CATCAAGTTGGCCCTTAACCACTAAGGATTTATCAAGATTTTTTGGCAGACTTATACATGTC	1017
Db	962	CATCAAGTTGGCCCTTAACCACTAAGGATTTATCAAGATTTTTTGGCAGACTTATACATGTC	1021
Qy	1018	TAGGATC 1024	
Db	1022	TAGGATC 1028	

## RESULT 2

US-09-397-787-68  
; Sequence 68, Application US/09397787  
; Patent No. 6468758

TELEPHONE NO. 0400750  
; GENERAL INFORMATION:

APPLICANT: Benson, Darin R.  
 APPLICANT: Lodes, Michael J.  
 APPLICANT: Mitcham, Jennifer L.  
 APPLICANT: King, Gordon E.  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN  
 TITILE OF INVENTION: CANCER THERAPY AND DIAGNOSIS  
 FILE REFERENCE: 210121.466C2  
 CURRENT APPLICATION NUMBER: US/09/397,787

; CURRENT AFFILIATION NUMBER: 05/03/3331,181  
 ; CURRENT FILING DATE: 1999-09-16  
 ; NUMBER OF SEQ ID NOS: 334

```

; NUMBER OF SEQ
; SOFTWARE: Fas
; SEQ ID NO 68
; LENGTH: 1664

```

### RESULT 3

US-09-397-787-267  
; Sequence 267, Application US/09397787  
; Patent No. 6468758

FACEBOOK NO. 0408738  
GENERAL INFORMATION:

; APPLICANT: Benson, Darin R.

US-09-397-787-68									
Query Match		80.9%; Score 828.4; DB 4; Length 1664;							
Best Local Similarity		98.7%; Pred. No. 5.1e-230;							
Matches 835; Conservative 0; Mismatches 11; Indels 0; Gaps 0;									
QY	179	AGGCTACCGCACTGCGCGCGGATTCATGNACTCCAGAGGGCGCTGTGAAGCTC	238						
DB	17	AGGCGCGGCAGTCGCGGGTAGCCCGGATTCATGAACCTCCAGAGGGCGCTGTGAAGCTC	76						
QY	239	CAGCAAAACAAGTATCCTATCTCTTTTTCGCAACCCATGAACTGCAATTTCTAGGTCCCA	298						
DB	77	CAGCAAAACAAGTATCCTATCTCTTTTTCGCAACCCATGAACTGCAATTTCTAGGTCCCA	136						
QY	299	AAGACCTTTTCCATATPAGGAGTACAAGACAAAGTTTGGAAAGTCAAAACAAACGGAAG	358						
DB	137	AAGACCTTTTCCATATPAGGAGTACAAGACAAAGTTTGGAAAGTCAAAACAAACGGAAG	196						
QY	359	GATTTACGAGGATTTCTGGGAAATAGAAAATACCCAGGAGTAAAGTTTACTTGGCTACC	418						
DB	197	GATTTACGAGGATTTCTGGGAAATAGAAAATACCCAGGAGTAAAGTTTACTTGGCTACC	256						
QY	419	AGGCAATTCAGCAACAGAGCTCTTCAGAAACTCAGGGAGAAAGTGGAAATACTTGCAGATG	478						
DB	257	AGGCAATTCAGCAACAGAGCTCTTCAGAAACTCAGGGAGAAAGTGGAAATACTTGCAGATG	316						
QY	479	CAAGCAGTAGGAGAAAGGTGATAGAGTAGAAGAGATGAAAAGGCAAAAGAAAGAAATG	538						
DB	317	CAAGCAGTAGGAGAAAGGTGATAGAGTAGAAGAGATGAAAAGGCAAAAGAAAGAAATG	376						
QY	539	AAAAGCAGGCTCAAAACGAAAAGTCATATCTTCAAGAAATCCTCTAAACAGTCCC	598						
DB	377	AAAAGCAGGCTCAAAACGAAAAGTCATATCTTCAAGAAATCCTCTAAACAGTCCC	436						
QY	599	GGAATCTTCAGGAGATGAAGATGCAAAAGACTGCAAAAGAGAGGAAAAACAAAGACAGCT	658						
DB	437	GGAATCTTCAGGAGATGAAGATGCAAAAGACTGCAAAAGAGAGGAAAAACAAAGACAGCT	496						
QY	659	CTGAGGTGAGATGCGGGCAACGACACAAAGAAACACAACTTCAGACTTCGAGAAACCA	718						
DB	497	CTGAGGTGAGATGCGGGCAACGACACAAAGAAACACAACTTCAGACTTCGAGAAACCA	556						
QY	719	GTGAAGGACCTAACCTACCATATGAATCTGATATTAAGAGAAACCAAGAGGTTA	778						
DB	557	GTGAAGGACCTAACCTACCATATGAATCTGATATTAAGAGAAACCAAGAGGTTA	616						
QY	779	TATGTTTGGTTGCTTAATATCTTTGGATTTGATATGAACCAACACATAGTCTTGTGTGC	838						
DB	617	TATGTTTGGTTGCTTAATATCTTTGGATTTGATATGAACCAACACATAGTCTTGTGTGC	676						
QY	839	ATTGACAGAACCCGAGTTGTATCATATTTCAATATTCCTCTGTGTGTGTGTGGGG	898						
DB	677	ATTGACAGAACCCGAGTTGTATGATCATATTTTCAATATTCCTCTGTGTGTGTGTGGGG	736						
QY	899	GGAAAACACATTTAGCCTTTTTTAAAGTTACTGATTAATTTCAATGTTATTTGGTTGC	958						
DB	737	GGAAAACACATTTAGCCTTTTTTAAAGTTACTGATTAATTTCAATGTTATTTGGTTGC	796						
QY	959	ATGAAGTTGCCCTTAACCACTAAGGATTTATCAAGATTTTTTGCAGACTTTATACATGCT	1018						
DB	797	ATGAAGTTGCCCTTAACCACTAAGGATTTATCAAGATTTTTTGCAGACTTTATACATGCT	856						
QY	1019	AGGATC 1024							
DB	857	AGGATC 862							

RESULT 3  
US-09-397-787-267  
; Sequence 267, Application US/09397787  
; Patent No. 6468758  
; GENERAL INFORMATION:  
; APPLICANT: Benson, Darin R.

```
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 267
; LENGTH: 1664
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-397-787-267

Query Match      80.9%; Score 828.4; DB 4; Length 1664;
Best Local Similarity 98.7%; Pred. No. 5.1e-230;
Matches 835; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 179 AGGGCTACCCGCACTGCGCGCGCGGATGATGAACCTCCAGAGGGGCGCTGTGAAGCCTC 238
DB 17 AGGGCGGCGCAGTCGCGGTAGCCGGATTGATGNACTCCAGAGGGGCGCTGTGAAGCCTC 76
QY 239 CAGCAAAACAAGTATCTTCTCTTTTGGCACCCCATGAACCTGCAATTTCTAGGTCCCA 298
DB 77 CAGCAAAACAAGTATCTTCTCTTTTGGCACCCCATGAACCTGCAATTTCTAGGTCCCA 136
QY 299 AAGACCTTTTCCATATAAGAGGTACAAAGACAAGTTTGGAAAGTCAAAACGGAAG 358
DB 137 AAGACCTTTTCCATATAAGAGGTACAAAGACAAGTTTGGAAAGTCAAAACGGAAG 196
QY 359 GATTTAAAGAGGATGTCGGAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTTACC 418
DB 197 GATTTAAAGAGGATGTCGGAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTTACC 256
QY 419 AGCAATTCAGCAACAGAGCTCTTTCAGAACTGAGGGAGAGGTGGAAATTAATCTGCAGATG 478
DB 257 AGCAATTCAGCAACAGAGCTCTTTCAGAACTGAGGGAGAGGTGGAAATTAATCTGCAGATG 316
QY 479 CAAGCAGTGAAGAGAGGTGATAGATAGAGNAGATGAAAAGGCAAAAGAAAGATG 538
DB 317 CAAGCAGTGAAGAGAGGTGATAGATAGAGNAGATGAAAAGGCAAAAGAAAGATG 376
QY 539 AAAAAAGCAGCTCAAAACGGAAGATCATATCTTCAAGAAATCTCTTAAACAGTCCC 598
DB 377 AAAAAAGCAGCTCAAAACGGAAGATCATATCTTCAAGAAATCTCTTAAACAGTCCC 436
QY 599 GGAATCTCCAGGAGTGAAGATGACAAAGATGCAAAAGAGGAGGAAACAAAGAGCAGCT 658
DB 437 GGAATCTCCAGGAGTGAAGATGACAAAGATGCAAAAGAGGAGGAAACAAAGAGCAGCT 496
QY 659 CTGAGGCTGAGATGCGGGCAACGACACAGAAACACAACTTCAGATTCGAGAAACCA 718
DB 497 CTGAGGCTGAGATGCGGGCAACGACACAGAAACACAACTTCAGATTCGAGAAACCA 556
QY 719 GTGAAGGAGCTTAACCTACCATATATGATGCTGCATATTAAGAGAAACCAAGAGGTTA 778
DB 557 GTGAAGGAGCTTAACCTACCATATATGATGCTGCATATTAAGAGAAACCAAGAGGTTA 616
QY 779 TATGTTTGGTGTCTAATATTTCTTGGATTTGATATGAACCAACACAPAGTCTTTGTGTC 838
DB 617 TATGTTTGGTGTCTAATATTTCTTGGATTTGATATGAACCAACACAPAGTCTTTGTGTC 676
QY 839 ATTGACAGAACCCAGTTGTATGATCATTAATTCATATTCCTCTCTGTTGTGTTGGGG 898
DB 677 ATTGACAGAACCCAGTTGTATGATCATTAATTCATATTCCTCTCTGTTGTGTTGGGG 736
QY 899 GGAAGAGACATTTAGCCTTTTAAAGTTACTGATTTAATTCATGTTATTTGTTGTC 958
DB 737 GGAAGAGACATTTAGCCTTTTAAAGTTACTGATTTAATTCATGTTATTTGTTGTC 796
QY 959 ATGAAGTTGCGCCTTAACCACTAAGGATTATCAAGATTTTTCGCGCAGACTTTATACATGCT 1018

Db 797 ATGAAGTTGCGCCTTAACCACTAAGGATTATCAAGATTTTTCGCGCAGACTTTATACATGCT 856
QY 1019 AGGATC 1024
DB 857 AGGATC 862.

RESULT 4
US-09-328-111-318/c
; Sequence 318, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 318
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(645)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-318

Query Match      46.0%; Score 471; DB 3; Length 645;
Best Local Similarity 95.8%; Pred. No. 1.1e-126;
Matches 525; Conservative 0; Mismatches 18; Indels 5; Gaps 4;

QY 320 AGTACAAAGACAAAGTTTGGAAAGTCAAAACGGAAGGATTTAAACGAAGGATTTGTGG 379
DB 557 AGGTCCAANGCCAGTTTGAAGAGTC-AACCAACGGAAGGATTTAAACGAAGGATTTGTGA 499
QY 380 AATAGAAAATTAACCCAGGATTAAGTTTACTGGCTTACAGGCAATTCAGCAACAGAGCT 439
DB 498 AATAGAAAAT--ACCCAGNAGTAAAGTTTACTGG-TACCAGGCAATTCAGCAACAGAGT 442
QY 440 CTTTCAGAACTGAGGAGAGGTGGAAATTAATCTGCAGATCAAGCAGTGAAGAGAGGTTG 499
DB 441 CTTTCAGAACTGA-GAGAGAGGTGGAAATTAATCTGCAGATCAAGCAGTGAAGAGAGGTTG 383
QY 500 ATAGAGTAGAAGAGTGGAAAGGCAAAAGAAAGAAATGAAAAGCAGGCTCAAAACGGA 559
DB 382 ATAGAGTAGAAGAGTGGAAAGGCAAAAGAAAGAAATGAAAAGCAGGCTCAAAACGGA 323
QY 560 AAAAGTCATATATTCTCAAGAAATCTCTTAAACAGTCCCGGAAATCTCCAGGAGATGAAG 619
DB 322 AAAAGTCATATATTCTCAAGAAATCTCTTAAACAGTCCCGGAAATCTCCAGGAGATGAAG 263
QY 620 ATGACAAAGACTGCAAGAGAGGAAACAAAGAGCAGCTCTGAGGCTGGAGATGCGGCA 679
DB 262 ATGACAAAGACTGCAAGAGAGGAAACAAAGAGCAGCTCTGAGGCTGGAGATGCGGCA 203
QY 680 ACGACAAAGAAACAACTTCAGACTTCGAGAAAACCAAGGAGGACCTTAACCTACCAT 739
```

Db	202	ACGCACACAGAAACACAACCTTCAGACTTTGCAGAAAAACCGTGAAGGACCTAACTACCAT	143
Qy	740	AATGAATGCTGCATATTAAAGAGAAACCAACAAGAGTTATATGTTTGGTTGCTCTAATATT	799
Db	142	AATGAATGCTGCATATTAAAGAGAAACCAACAAGAGTTATATGTTTGGTTGCTCTAATATT	83
Qy	800	CTTGGAATTGGATGATGAACCAACACATAGTCTCTTGTTCATTGACAGAACCCAGTTTGT	859
Db	82	CTTGGAATTGGATGATGAACCAACACATAGTCTCTTGTTCATTGACAGAACCCAGTTTGT	23
Qy	860	ATGTACAT	867
Db	22	ATGTACCT	15

RESULT 5  
US-08-760-745-4  
; Sequence 4, Application US/08760745  
; Patent No. 5972658  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Goli, Surya K.  
; APPLICANT: Murry, Lynn E.  
; TITLE OF INVENTION: LUNG GROWTH FACTOR VARIANT  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

Query Match	18.9%;	Score 193.4;	DB 2;	Length 2376;
Best Local Similarity	62.5%;	Pred. No. 7.1e-46;		
Matches 302;	Conservative	0;	Mismatches 181;	Indels 0;
Gaps	0;			

  

QY	23	GGCCCGCGGGAGCGCGCAATTCTGTCGGCCCCGGGGGGGGCGCTCTCCCGCATCTTC	82
Db	221	GGCCCGAGCGAGCGCGCCGCCACCGCGCACGCGCAAACTTTGGGCTCGCGCTTCCCGG	280
QY	83	CGCGCGCACCAAGACTACACAGAAAGGAGCGGGCTGGGATGGCGCGTTCGCGGCCCCCGG	142

281	CCCGCGCGGAGCCCGGGGCGCCCGGAGCCCGCCATGTCGCGATCCAAACCGCGAAGG	340
143	AGTACAAAGCGGCGACCTGCTCTTCGCCAAGATGAAGGGCTACCCGCACTGGCGGGCC	202
341	AGTACAAATCCGGGACCTGGTGTTCGCAAGATGAAGGGCTACCCACACTGGCGGGCC	400
203	GGATTGATNAATCCCAAGAGGGCGCTGTGAAGCCTCCAGCAAAACAAGTATCTATCTTCT	262
401	GGATTGACGAGATGCTTGAGGCTGCCGTGAATCAACAGCCAAATACCAAGTCTTTT	460
263	TTTTTGGCACCCATGAACCTGCATTTCTAGTGTCCAAAGACCTTTTCCATATAGGAGT	322
461	TTTTTCGGGACCCACGACGCGCATTCCTGGGCCCCAAAGACTCTTCCCTTACGAGGAAT	520
323	ACAAAGACAGTTTGGNAAGTCAACAAACGGAAGGATTAAACGAGGATTGTGGGAA	382
521	CCAAGGAGAAGTTTGGCAGCCCCAACAGAGGAAGGGTTTCAGCGAGGGCTGTGGGAGA	580
383	TAGAAATAACCCAGGAGTAAAGTTTACTCGCTACCAAGCAATTCAGCAACAGAGCTCTT	442
581	TCGAGAACAAACCTTACTGTCAAGGCTTCGGGCTATCAGTCTCCAGAAAAGAGCTGTG	640
443	CAGAACTCAGGGGAGAGGTGGAATACTGCAGATGCAAGCAGTGCAGGAAGAAGTGATA	502
641	TGGAAGGCCCTGAACCAAGACCCCGAAGCTGCAGAGGGGTGACGCTGATAGAAGGGGAATG	700
503	GAG 505	
701	CAG 703	

RESULT 6  
US-08-760-745-2  
; Sequence 2, Application US/08760745  
; Patent No. 5972658  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Goli, Surya K.  
; APPLICANT: Murry, Lynn E.  
; TITLE OF INVENTION: LUNG GROWTH FACTOR VARIANT  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/760,745  
; FILING DATE: Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PP-0169 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 869 base pairs  
; TYPE: nucleic acid

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2004, 10:16:26 ; Search time 395 Seconds

(without alignments)  
9137.614 Million cell updates/sec

Title: US-09-787-328B-3

Perfect score: 1024

Sequence: 1 accgctcgtccgcccgctt.....acttatacatgtctaggatc 1024

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1000	97.7	1898	9	US-09-768-826-22
2	985.6	96.2	1919	9	US-09-925-301-145
3	980	95.7	3805	13	US-10-247-671-8
4	828.4	80.9	1664	9	US-09-876-889-68
5	828.4	80.9	1664	9	US-09-876-889-267
6	778	76.0	2920	11	US-09-866-050A-541
7	778	76.0	2920	15	US-10-152-661-541
8	471	46.0	645	10	US-09-879-536-318
9	248.6	24.3	398	10	US-09-960-352-8612
10	239.6	23.4	288	10	US-09-960-352-1025
11	193.4	18.9	2376	9	US-09-938-985A-4
12	193.4	18.9	2376	15	US-10-207-791-3
13	189	18.5	723	15	US-10-207-791-1
14	189	18.5	2133	9	US-09-925-302-68
15	188.6	18.4	869	9	US-09-938-985A-2

16	173	16.9	3316	13	US-10-133-013-166	Sequence 166, App
17	172.6	16.9	800	9	US-09-910-943-585	Sequence 585, App
18	168.8	16.5	480	11	US-09-918-995-32444	Sequence 32444, A
19	159.8	15.6	1297	9	US-09-768-826-24	Sequence 24, Appl
20	159.8	15.6	2031	11	US-09-823-187-23	Sequence 23, Appl
21	159.8	15.6	2272	11	US-09-946-374-307	Sequence 307, App
22	159.8	15.6	2272	12	US-10-011-795A-307	Sequence 307, App
23	159.8	15.6	2272	13	US-10-015-387A-307	Sequence 307, App
24	159.8	15.6	2272	13	US-10-006-130A-307	Sequence 307, App
25	159.8	15.6	2272	13	US-10-199-672-345	Sequence 345, App
26	159.8	15.6	2272	13	US-10-006-172A-307	Sequence 307, App
27	159.8	15.6	2272	13	US-10-187-749-345	Sequence 345, App
28	159.8	15.6	2272	13	US-10-194-457-345	Sequence 345, App
29	159.8	15.6	2272	13	US-10-184-642-345	Sequence 345, App
30	159.8	15.6	2272	13	US-10-196-747-345	Sequence 345, App
31	159.8	15.6	2272	13	US-10-015-392A-307	Sequence 307, App
32	159.8	15.6	2272	13	US-10-017-253A-307	Sequence 307, App
33	159.8	15.6	2272	13	US-10-173-689-345	Sequence 345, App
34	159.8	15.6	2272	13	US-10-173-690-345	Sequence 345, App
35	159.8	15.6	2272	13	US-10-173-691-345	Sequence 345, App
36	159.8	15.6	2272	13	US-10-173-692-345	Sequence 345, App
37	159.8	15.6	2272	13	US-10-173-694-345	Sequence 345, App
38	159.8	15.6	2272	13	US-10-173-698-345	Sequence 345, App
39	159.8	15.6	2272	13	US-10-173-699-345	Sequence 345, App
40	159.8	15.6	2272	13	US-10-173-707-345	Sequence 345, App
41	159.8	15.6	2272	13	US-10-174-569-345	Sequence 345, App
42	159.8	15.6	2272	13	US-10-174-583-345	Sequence 345, App
43	159.8	15.6	2272	13	US-10-174-587-345	Sequence 345, App
44	159.8	15.6	2272	13	US-10-174-589-345	Sequence 345, App
45	159.8	15.6	2272	13	US-10-174-591-345	Sequence 345, App

ALIGNMENTS

RESULT 1

US-09-768-826-22  
; Sequence 22, Application US/09768826  
; Patent No. US20020012966A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi et al.  
; TITLE OF INVENTION: 18 human secreted proteins  
; FILE REFERENCE: PFS12P1  
; CURRENT APPLICATION NUMBER: US/09/768,826  
; CURRENT FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: PCT/US00/22350  
; PRIOR FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: 60/148,759  
; PRIOR FILING DATE: 1999-08-16  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 1898  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-768-826-22

Query Match	97.7%	Score	1000	DB	9	Length	1898	
Best Local Similarity	99.8%	Pred. No.	7e-267					
Matches 1022	Conservative	0	Mismatches	0	Indels	2	Gaps	2
QY	2	CGCTCGTCCGCCCGCTTTAGCGCCGCGGGAGCGC-GCGCAATTCGTCCGCCCGCGGG	60					
Db	30	CGCTCGTCCGCCCGCTTTAGCGCCGCGGGAGCGCGCGCGAATTCGTCCGCCCGCGGG	89					
QY	61	GGGCGCGCTCCCGGCATCTTCGCGCGACCAAGGACTACAGGAAGGGAGCGGCTCGG	120					
Db	90	GGGCGCGCTCCCGGCATCTTCGCGCGACCAAGGACTACAGGAAGGGAGCGGCTCGG	149					
QY	121	ATGGCGCGTCCGGGCCCCCGGAGTACAAAGCGGGGACCTGGTCTTCGCCAAGATGAAG	180					
Db	150	ATGGCGCGTCCGGGCCCCCGGAGTACAAAGCGGGGCGACTGGTCTTCGCCAAGATGAAG	209					

181 GGCTACCCGCACTGGCGCGCGGATGTGATGAACCTCCAGAGGGCGCTGTGAAGCTCCA 240  
Db  
210 GGCTACCCGCACTGGCGCGCGGATGTGATGAACCTCCAGAGGGCGCTGTGAAGCTCCA 269  
Qy 241 GCAACACAGTATCCTATCTCTTTTGGACCCATGAACCTGCAATTTCTAGTCCCAA 300  
Db 270 GCAACACAGTATCCTATCTCTTTTGGACCCATGAACCTGCAATTTCTAGTCCCAA 329  
Qy 301 GACCTTTTCCATATAAGAGGTACAAAGACAAGTTTGGAAAGTCAAAACGGAAGGA 360  
Db 330 GACCTTTTCCATATAAGAGGTACAAAGACAAGTTTGGAAAGTCAAAACGGAAGGA 389  
Qy 361 TTTAAGCAAGAGTGTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACCAG 420  
Db 390 TTTAAGCAAGAGTGTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACCAG 449  
Qy 421 GCAATTCAGCAACAGAGCTCTTCAGAACTGAGGGAGAAAGGTGGGAAATACTCAGATGA 480  
Db 450 GCAATTCAGCAACAGAGCTCTTCAGAACTGAGGGAGAAAGGTGGGAAATACTCAGATGA 509  
Qy 481 AGCAGTGAAGAGAGGTGATAGAGTGAAGAGAGTGAAGAGGCAAAAGGAAGATGA 540  
Db 510 AGCAGTGAAGAGAGGTGATAGAGTGAAGAGAGTGAAGAGGCAAAAGGAAGATGA 569  
Qy 541 AAAGCAGGCTCAAAACGGAAGAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCCGG 600  
Db 570 AAAGCAGGCTCAAAACGGAAGAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCCGG 629  
Qy 601 AAATCTCAGGAGATGAAGATGACAAAGACTGCAAGAGAGAGGAAACAAAGAGAGCTCT 660  
Db 630 AAATCTCAGGAGATGAAGATGACAAAGACTGCAAGAGAGAGGAAACAAAGAGAGCTCT 689  
Qy 661 GAGGTGAGAGTGGCGGCAACGACACAAAGACAACTTCAGACTTCGAGAAACACAGT 720  
Db 690 GAGGTGAGAGTGGCGGCAACGACACAAAGACAACTTCAGACTTCGAGAAACACAGT 749  
Qy 721 GAAGGACCTAACTACCAATATGAATGCTGCATATTAAGAGAAACCAAGAGAGTTATA 780  
Db 750 GAAGGACCTAACTACCAATATGAATGCTGCATATTAAGAGAAACCAAGAGAGTTATA 809  
Qy 781 TGTGTTGTTGCTAATATTTCTTGGATTTGATATGAACCAACATAGTCTCTTGTGTCAT 840  
Db 810 TGTGTTGTTGCTAATATTTCTTGGATTTGATATGAACCAACATAGTCTCTTGTGTCAT 869  
Qy 841 TCACAGAACCCAGTTGATGATGATATTAATTCATATCTCTCTGTTGTTTCGGGGG 900  
Db 870 TCACAGAACCCAGTTGATGATGATATTAATTCATATCTCTCTGTTGTTTCGGGGG 929  
Qy 901 AAAAGACATTTTAGCCTTTTAAAGTTTACTGATTTAATTTTCAATGTTTGTGTCAT 960  
Db 930 AAAAGACATTTTAGCCTTTTAAAGTTTACTGATTTAATTTTCAATGTTTGTGTCAT 989  
Qy 961 GAAGTTGCCCTTAACCACTAAGATTAACAGATTTTGGCGAGACTTATACATGCTAG 1020  
Db 990 GAAGTTGCCCTT-ACCACCTAAGATTAACAGATTTTGGCGAGACTTATACATGCTAG 1048  
Qy 1021 GATC 1024  
Db 1049 GATC 1052

RESULT 2  
US-09-925-301-145  
; Sequence 145, Application US/09925301  
; Patent No. US20020052308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA106  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05882  
; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1694  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 145  
; LENGTH: 1919  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1882)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (1898)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: misc\_feature  
; LOCATION: (1919)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-301-145  
Query Match 96.2%; Score 985.6; DB 9; Length 1919;  
Best Local Similarity 99.4%; Pred. No. 7.1e-263;  
Matches 1018; Conservative 2; Mismatches 1; Indels 3; Gaps 3;  
Qy 2 CGGCTCGTCCCGCGGCTTTGAGGCGCGCGGAGCGC-GCGCAATTCGTGCGCCCGCGG 60  
Db 15 CGGCTCGTCCCGCGGCTTTGAGGCGCGCGGAGCGCGCAATTCGTGCGCCCGCGG 74  
Qy 61 GGGCGGCTCTCCCGCATCTTCGGCGGCGACCAAGACTTACCAAGAGGGAGCGCTGGG 120  
Db 75 GGGCGGCTCTCCCGCATCTTCGGCGGCGACCAAGACTTACCAGGAAGGGAGCGCTGGG 134  
Qy 121 ATGGCGGCTCGCGCGCGCGAGGTACAAAGCGGCGACCTGCTTCGCCAAGATGAAG 180  
Db 135 ATGGCGGCTCGG-GGCGCGGAGTACAAAGCGGCGACCTGCTTCGCCAAGATGAAG 193  
Qy 181 GGCTACCCGCACTGGCGCGCGCGGATTTGATGAACCTCCAGAGGGCGCTGTGAAGCTCCA 240  
Db 194 GGCTACCCGCACTGGCGCGCGCGGATTTGATGAACCTCCAGAG-GGCGCTGTGAAGCTTCA 252  
Qy 241 GCAACACAGTATCTCTCTTTTGGCAACCAATGAACCTGCAATTTCTAGGTCCCAA 300  
Db 253 GCAACACAGTATCTCTCTTTTGGCAACCAATGAACCTGCAATTTCTAGGTCCCAA 312  
Qy 301 GACCTTTTCCATATAAGAGGTACAAAGACAAGTTTGGAAAGTCAAAACGGAAGGA 360  
Db 313 GACCTTTTCCATATAAGAGGTACAAAGACAAGTTTGGAAAGTCAAAACGGAAGGA 372  
Qy 361 TTTAACGAAGAGTTGTGGGAAATAGAAATAACCCAGAGTAAAGTTTACTGGCTACCAG 420  
Db 373 TTTAACGAAGAGTTGTGGGAAATAGAAATAACCCAGGAGTAAAGTTTACTGGCTACCAG 432  
Qy 421 GCAATTCAGCAACAGAGCTCTTCAGAACTGAGGGAGAGGTGGAAATCTGCGAGATGA 480  
Db 433 GCAATTCAGCAACAGAGCTCTTCAGAACTGAGGGAGAGGTGGAAATCTGCGAGATGA 492  
Qy 481 AGCAGTGAAGAGAGGTGATAGAGTGAAGAGTGAAGAGGCAAAAGGAAGATGA 540  
Db 493 AGCAGTGAAGAGAGGTGATAGAGTGAAGAGTGAAGAGGCAAAAGGAAGATGA 552  
Qy 541 AAAGCAGGCTCAAAACGGAAGAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCCGG 600  
Db 553 AAAGCAGGCTCAAAACGGAAGAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCCGG 612  
Qy 601 AAATCTCAGGAGATGAAGATGACAAAGCTTGCAGAGAGGAGGAAACAAAGAGAGCTCT 660  
Db 613 AAATCTCAGGAGATGAAGATGACAAAGCTTGCAGAGAGGAGGAAACAAAGAGAGCTCT 672  
Qy 661 GAGGTGAGAGTGGGCAACGACACAGAAACCAACTTCAGACTTCAGAAACACAGT 720  
Db 673 GAGGTGAGAGTGGGCAACGACACAGAAACCAACTTCAGACTTCAGAAACACAGT 732  
Qy 721 GAAGGACCTAACTACCAATATGAATGCTGCATATTAAGAGAAACCAAGAGAGTTATA 780



Db 733 GAAGGGACCTAACTACCATAATGCTGCATATTAGAGAAACCAAGAGTTATA 792  
Qy 781 TGTTGGTGTCTAAATATCTTGGATTGATATGAACCAACACATAGTCTCTTGTGTCAT 840  
Db 793 TGTTGGTGTCTAAATATCTTGGATTGATATGAACCAACACATAGTCTCTTGTGTCAT 852  
Qy 841 TGACAGAACCCAGTTTGTATGTACATATTTCAATATCTCTCTGTGTTGTGTCGGGGG 900  
Db 853 TGACAGAACCCAGTTTGTATGTACATATTTCAATATCTCTCTGTGTTGTGTCGGGGG 912  
Qy 901 AAAAGACATTTAGCCCTTTTAAAGTTACTGATTTAAATTCATGTTATTTGGTTCAT 960  
Db 913 AAAAGACATTTAGCCCTTTTAAAGTTACTGATTTAAATTCATGTTATTTGGTTCAT 972  
Qy 961 GAAGTTGCCCTTAACCACTAAGGATTTCAAGATTTTGGCGACCTTATACATGTCCTAG 1020  
Db 973 GAAGTTGCCCTTAACCACTAAGGATTTCAAGATTTTGGCGACCTTATACATGTCCTAG 1032  
Qy 1021 GATC 1024  
Db 1033 GATC 1036

## RESULT 3

US-10-247-671-8  
; Sequence 8, Application US/10247671  
; Publication No. US20030194721A1  
; GENERAL INFORMATION:  
; APPLICANT: Mikita, Thomas  
; APPLICANT: Shiffman, Dov  
; APPLICANT: Porter, Gordon, J.  
; APPLICANT: Kaser, Matthew R.  
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS  
; FILE REFERENCE: PA-0050 US  
; CURRENT APPLICATION NUMBER: US/10/247,671  
; CURRENT FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: 60/323,784  
; NUMBER OF SEQ ID NOS: 186  
; SOFTWARE: PERL Program  
; SEQ ID NO 8  
; LENGTH: 3805  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. US20030194721A1 2676869CB1  
; NAME/KEY: unsure  
; LOCATION: 3788  
; OTHER INFORMATION: a, t, c, g, or other  
US-10-247-671-8

Query Match 95.7%; Score 980; DB 13; Length 3805;

Best Local Similarity 99.1%; Pred. No. 4e-261;

Matches 1017; Conservative 0; Mismatches 5; Indels 4; Gaps 3;

Qy 1 ACCGCTCGTCGCGCGCTTAGCGCCCGCGGGAGCGCGGCAATTCGTGCGCCCGCGGG 60  
Db 4 ACCGCTCGTCG-CGCGCTTAGCGCCCGCGGGAG-CGCGGCAATTCGTGCGCCCGCGGG 61  
Qy 61 GGGCGCGCC--TCGCGCATCTTCGCGCGACCAAGGACTACCAAGGAGGGAGCGGCTG 118  
Db 62 GGGCGCGCCATCCGGCAGTCTTCGCGCGACCAAGGACTACCAAGGAGGGAGCGGCTG 121  
Qy 119 GGATGCGCGCTCCGCGCCCGCGAGTACAAAGCGGCGACCTCGTCTTCGCCAAGATGA 178  
Db 122 GGATGCGCGCTCCGCGCCCGCGAGTACAAAGCGGCGACCTCGTCTTCGCCAAGATGA 181  
Qy 179 AGGCTTACCGGACTGCGCGCGCGCGGATGATGAATCCAGAGGGCGCTGTGAAGCTC 238  
Db 182 AGGCTTACCGGACTGCGCGCGCGCGGATGATGAATCCAGAGGGCGCTGTGAAGCTC 241

Qy 239 CAGCAACCAAGTATCCTATCTCTTTTGGCAACCCATGAAATGCAATTTCTAGGTCCCA 298  
Db 242 CAGCAACCAAGTATCCTATCTCTTTTGGCAACCCATGAAATGCAATTTCTAGGTCCCA 301  
Qy 299 AAGACCTTTTCCATATTAAGGAGTACAAGACAAGTTTGGAAAGTCAAAACAACGGAAG 358  
Db 302 AAGACCTTTTCCATATTAAGGAGTACAAGACAAGTTTGGAAAGTCAAAACAACGGAAG 361  
Qy 359 GATTTAACGAAGGATTTGTGGAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACC 418  
Db 362 GATTTAACGAAGGATTTGTGGAATAGAAATTAACCCAGGAGTAAAGTTTACTGGCTACC 421  
Qy 419 AGGCAATTCAGCAACAGAGCTCTTTCAGAACTCAGGGAGAGGTGGAATATCTGCAGATG 478  
Db 422 AGGCAATTCAGCAACAGAGCTCTTTCAGAACTCAGGGAGAGGTGGAATATCTGCAGATG 481  
Qy 479 CAAGCAGTAGGGAAGAGGTGATAGATGAGAAGATGGAAGAGGCAAGAAAGAAATG 538  
Db 482 CAAGCAGTAGGGAAGAGGTGATAGATGAGAAGATGGAAGAGGCAAGAAAGAAATG 541  
Qy 539 AAAAAGCAGGCTCAAAACGGAAGAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCC 598  
Db 542 AAAAAGCAGGCTCAAAACGGAAGAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCC 601  
Qy 599 GGAATCTCCAGGAGATGAAGATGACAAAGACTGCAAAAGAGAGGAAACAAAGAGCAGCT 658  
Db 602 GGAATCTCCAGGAGATGAAGATGACAAAGACTGCAAAAGAGAGGAAACAAAGAGCAGCT 661  
Qy 659 CTGAGGGTGAGATGCGGGCAACGACACAAGAAACAACACTTCAGACTTCGAGAAAACCA 718  
Db 662 CTGAGGGTGAGATGCGGGCAACGACACAAGAAACAACACTTCAGACTTCGAGAAAACCA 721  
Qy 719 GTCAAGGGACCTTAACCTACCATATGATGCTCATATTAAGAGAAACCAAGAAAGTTA 778  
Db 722 GTCAAGGGACCTTAACCTACCATATGATGCTCATATTAAGAGAAACCAAGAAAGTTA 781  
Qy 779 TATGTTTGGTGTCTAATATCTTGGATTGATGAACCAACACATAGTCTTGTGTC 838  
Db 782 TATGTTTGGTGTCTAATATCTTGGATTGATGAACCAACACATAGTCTTGTGTC 841  
Qy 839 ATTGACAGAACCCAGTTTGTATGTACATATTCATATTCCTCTCTGTTGTGTTGGGG 898  
Db 842 ATTGACAGAACCCAGTTTGTATGTACATATTCATATTCCTCTCTGTTGTGTTGGGG 901  
Qy 899 GGAAGAGACATTTAGCCCTTTTAAAGTTACTGATTTAAATTTCAATTTGTTGTTGTC 958  
Db 902 GGAAGAGACATTTAGCCCTTTTAAAGTTACTGATTTAAATTTCAATTTGTTGTTGTC 961  
Qy 959 ATGAAGTTGCCCTTAACCACTAAGGATTTCAAGATTTTGGCAGACTTATACATGCTCT 1018  
Db 962 ATGAAGTTGCCCTTAACCACTAAGGATTTCAAGATTTTGGCAGACTTATACATGCTCT 1021  
Qy 1019 AGGATC 1024  
Db 1022 AGGATC 1027

## RESULT 4

US-09-876-889-68  
; Sequence 68, Application US/09876889  
; Patent No. US20020076715A1  
; GENERAL INFORMATION:  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: King, Gordon E.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN  
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS  
; FILE REFERENCE: 210121.466C3  
; CURRENT APPLICATION NUMBER: US/09/876,889  
; CURRENT FILING DATE: 2001-06-06  
; NUMBER OF SEQ ID NOS: 353

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 68
; LENGTH: 1664
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-876-889-68

Query Match      80.9%; Score 828.4; DB 9; Length 1664;
Best Local Similarity 98.7%; Pred. No. 3.1e-219;
Matches 835; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 179 AGGGGTACCCGCACTGGCCGCGGATGATGAATCCAGAGGCGCTGTGAAGCCTC 238
DB 17 AGGGCGGCGACGTGCGGGTAGCCGGATTGATGAATCCAGAGGCGCTGTGAAGCCTC 76
QY 239 CAGCAAAACAAGTATCTATCTCTTTTGGACCCATGAATCTAGTCTCCCA 298
DB 77 CAGCAAAACAAGTATCTATCTCTTTTGGACCCATGAATCTAGTCTCCCA 136
QY 299 AAGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAACAAACGGAAG 358
DB 137 AAGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAACAAACGGAAG 196
QY 359 GATTTAAACGAAGATTGTGGGAAATAGAAAATAAACCCAGGAGTAAAGTTTACTGGCTACC 418
DB 197 GATTTAAACGAAGATTGTGGGAAATAGAAAATAAACCCAGGAGTAAAGTTTACTGGCTACC 256
QY 419 AGGCAATTTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAAAGTGAATATCTGCAGATG 478
DB 257 AGGCAATTTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAAAGTGAATATCTGCAGATG 316
QY 479 CAAGCAGTGAGGAAGAGTGATAGTAGAAGAGTGAAGAGAGTGAAGAGCAAGAAAGATG 538
DB 317 CAAGCAGTGAGGAAGAGTGATAGTAGAAGAGTGAAGAGAGTGAAGAGCAAGAAAGATG 376
QY 539 AAAAAGCAGGCTCAAAACGGAAGAGTCAATATCTCTCAAGAAATCTCTCAACAGTCCC 598
DB 377 AAAAAGCAGGCTCAAAACGGAAGAGTCAATATCTCTCAAGAAATCTCTCAACAGTCCC 436
QY 599 GGAATCTTCAGGAGATGAAGATGACAAAGATGCAAGAGAGAGAAACAAACAGAGT 658
DB 437 GGAATCTTCAGGAGATGAAGATGACAAAGATGCAAGAGAGAGAAACAAACAGAGT 496
QY 659 CTGAGGTGGAGATGCGGCAACGACACAAAGAAACACAACTTCAGACTTCGAGAAACCA 718
DB 497 CTGAGGTGGAGATGCGGCAACGACACAAAGAAACACAACTTCAGACTTCGAGAAACCA 556
QY 719 GTGAAGGACCTTAACCTAATGCTGCTATTAAGAGAAACCAACAGAGAGTTA 778
DB 557 GTGAAGGACCTTAACCTAATGCTGCTATTAAGAGAAACCAACAGAGAGTTA 616
QY 779 TATGTTGGTGTCTAAATCTCTGGATTTGATATGAACCAACATAGTCTCTGTTGTC 838
DB 617 TATGTTGGTGTCTAAATCTCTGGATTTGATATGAACCAACATAGTCTCTGTTGTC 676
QY 839 ATTGACAGAACCCAGTTTGTATGATATTAATTCCTCTCTGTTGTTGTTGCGGG 898
DB 677 ATTGACAGAACCCAGTTTGTATGATATTAATTCCTCTCTGTTGTTGTTGCGGG 736
QY 899 GGAAAGACATTTTACCTTTTAAAGTTACTGATTTAAATTTTCAATGTTTATTTGTTGTC 958
DB 737 GGAAAGACATTTTACCTTTTAAAGTTACTGATTTAAATTTTCAATGTTTATTTGTTGTC 796
QY 959 ATGAAGTTCCTTAAACCACTAAGGATTAATCAAGATTTTGGCAGACTTATACATGCT 1018
DB 797 ATGAAGTTCCTTAAACCACTAAGGATTAATCAAGATTTTGGCAGACTTATACATGCT 856
QY 1019 AGGATC 1024
DB 857 AGGATC 862
```

RESULT 5

```
US-09-876-889-267
; Sequence 267, Application US/09876889
; Patent No. US20020076715A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; FILE REFERENCE: 210121.466C3
; CURRENT APPLICATION NUMBER: US/09/876.889
; CURRENT FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 353
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 267
; LENGTH: 1664
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-876-889-267
```

```
Query Match      80.9%; Score 828.4; DB 9; Length 1664;
Best Local Similarity 98.7%; Pred. No. 3.1e-219;
Matches 835; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 179 AGGGCTACCCGCACTGGCCGCGGATGATGAATCCAGAGGCGCTGTGAAGCCTC 238
DB 17 AGGGCGGCGACGTGCGGGTAGCCGGATTGATGAATCCAGAGGCGCTGTGAAGCCTC 76
QY 239 CAGCAAAACAAGTATCTATCTCTTTTGGACCCATGAATCTAGTCTCCCA 298
DB 77 CAGCAAAACAAGTATCTATCTCTTTTGGACCCATGAATCTAGTCTCCCA 136
QY 299 AAGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAACAAACGGAAG 358
DB 137 AAGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAACAAACGGAAG 196
QY 359 GATTTAAACGAAGATTGTGGGAAATAGAAAATAAACCCAGGAGTAAAGTTTACTGGCTACC 418
DB 197 GATTTAAACGAAGATTGTGGGAAATAGAAAATAAACCCAGGAGTAAAGTTTACTGGCTACC 256
QY 419 AGGCAATTTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAAAGTGAATATCTGCAGATG 478
DB 257 AGGCAATTTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAAAGTGAATATCTGCAGATG 316
QY 479 CAAGCAGTGAGGAAGAGTGATAGTAGAAGAGTGAAGAGAGTGAAGAGCAAGAAAGATG 538
DB 317 CAAGCAGTGAGGAAGAGTGATAGTAGAAGAGTGAAGAGAGTGAAGAGCAAGAAAGATG 376
QY 539 AAAAAGCAGGCTCAAAACGGAAGAGTCAATATCTCTCAAGAAATCTCTCAACAGTCCC 598
DB 377 AAAAAGCAGGCTCAAAACGGAAGAGTCAATATCTCTCAAGAAATCTCTCAACAGTCCC 436
QY 599 GGAATCTTCAGGAGATGAAGATGACAAAGATGCAAGAGAGAGAAACAAACAGAGT 658
DB 437 GGAATCTTCAGGAGATGAAGATGACAAAGATGCAAGAGAGAGAAACAAACAGAGT 496
QY 659 CTGAGGTGGAGATGCGGCGCAACGACACAAAGAAACACAACTTCAGACTTCGAGAAACCA 718
DB 497 CTGAGGTGGAGATGCGGCGCAACGACACAAAGAAACACAACTTCAGACTTCGAGAAACCA 556
QY 719 GTGAAGGACCTTAACCTAATGCTGCTATTAAGAGAAACCAACAGAGAGTTA 778
DB 557 GTGAAGGACCTTAACCTAATGCTGCTATTAAGAGAAACCAACAGAGAGTTA 616
QY 779 TATGTTGGTGTCTAAATCTCTGGATTTGATATGAACCAACATAGTCTCTGTTGTC 838
DB 617 TATGTTGGTGTCTAAATCTCTGGATTTGATATGAACCAACATAGTCTCTGTTGTC 676
QY 839 ATTGACAGAACCCAGTTTGTATGATATTAATTCCTCTCTGTTGTTGTTGCGGG 898
DB 677 ATTGACAGAACCCAGTTTGTATGATATTAATTCCTCTCTGTTGTTGTTGCGGG 736
```

Qy 899 GGAAGAGACATTTTAGCCCTTTTAAAGATTACTGATTAATTAATTCATGTTATTTGGTGC 958  
Db 737 GGAAGAGACATTTTAGCCCTTTTAAAGATTACTGATTAATTAATTCATGTTATTTGGTGC 796  
Qy 959 ATGAAGTTGCCCTTAACCACTAAGATTAATCAAGATTTTTCGCGAGACTTATACATGCT 1018  
Db 797 ATGAAGTTGCCCTTAACCACTAAGATTAATCAAGATTTTTCGCGAGACTTATACATGCT 856  
Qy 1019 AGGATC 1024  
Db 857 AGGATC 862  
RESULT 6  
US-09-866-050A-541  
; Sequence 541, Application US/09866050A  
; Publication No. US20030040471A1  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Murison, James G.  
; APPLICANT: Kumble, Krishanand D.  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; FILE REFERENCE: 11000.1011c4U  
; CURRENT APPLICATION NUMBER: US/09/866,050A  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 725  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 541  
; LENGTH: 2920  
; TYPE: DNA  
; ORGANISM: Rat  
US-09-866-050A-541

Query Match 76.0%; Score 778; DB 11; Length 2920;  
Best Local Similarity 89.1%; Pred. No. 4.6e-205;  
Matches 912; Conservative 0; Mismatches 95; Indels 17; Gaps 6;  
Qy 2 CCGCTCGTCCCGCCGCTTGAAGCCCGCGGGAGCGC-GCGCAATTCGTGCGCCCGCGGG 60  
Db 164 CCGCTCGAGCCCGCGCTCGAGGCGCGCGGAAGCGCGCGCTCGCTCGCTCCCGGGA 223  
Qy 61 GGGCGGCGCTCCCGCATCTTCGCGCGGACCAAGAGCTACCAAGAGGAGCGCTGGG 120  
Db 224 GGGCGGCGCTCCCGCATCTTCGCGCGGACCAAGAGCGAGCGCGGAGCGGAAACCAACGGG 283  
Qy 121 ATGGCGGCTCGCGCCCGCGAGTACAAAGCGCGGCGACCTGGTCTTCGCCAAGATGAAG 180  
Db 284 ATGGCGGCTCGCGCCCGCGAGTACAAAGCGGAGACCTGGTCTTCGCCAAGATGAAG 343  
Qy 181 GGCTACCGCGACTCGCGCGCGCGGATGATGAATCCCGAGAGCGCTGTGAAGCCTCCA 240  
Db 344 GGCTACCGCGACTCGCGCGCGCGGATGATGAATCCCGAGAGCGCGCTGAAGCCTCCA 403  
Qy 241 GCAACAGATCTCTATCTCTTTTGGCACCCATGAACTGCAATCTTAGTCCCAAA 300  
Db 404 GCAACAGATCTCTATCTCTTTTGGCACCCATGAACTGCAATCTTAGTCCCTAAA 463  
Qy 301 GACCTTTTCCATATAAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACAAACGGAAGGA 360  
Db 464 GACCTTTTCCATATAAGGATACAAAGACAAAGTTTGGAAAGTCAAAACAAACGGAAGGA 523  
Qy 361 TTTAACGAAGATTGTGGAAATAGAAAATACCAAGAGTAAAGTTTACTGGCTACCAAG 420  
Db 524 TTTAATGAAGATTATGGAAATTTGAAAATATCCAGGAGTAAATTTACTGGGTACCAAG 583  
Qy 421 GCATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGTGGAATATCTGCAAGTGA 480  
Db 584 ACAATTGAGCAACAGAGCTCTTCAGAACTGAGGAGAGTGGAATATCTGCAAGTGA 643

Qy 481 AGCAGTGAGGAAGAGGTGATAGAGTAGAAGATGGAAGGCAAGAAAGAAATGAA 540  
Db 644 AGCAGTGAGGAAGAGGTGACAGAGT---AGAAGATGGAAGGCAAGAAAGAAATGAA 700  
Qy 541 AAAGCAGGCTCAAAACGGAAGAGTCAATATCTTCAAGAAATCCTCTAAACAGTCCCGG 600  
Db 701 AAAGGAGGCTCAAAACGGAAGAGTCTACACTTCAAGAAAGTCTTCTAAACAGTCCCGG 760  
Qy 601 AAATCTCCAGGATGAAGATGACAAAGAGCTGCAAGAGAGGAGGAAACAAAGCAGCTCT 660  
Db 761 AAATCTCCAGGAGAGAGATGACAAAGAGCTGCAAGAGAGGAGGAGGAAACAAAGCAGCTCT 820  
Qy 661 GAGCGTGGAGATCGGCGCAACGACACAAAGAAACACAACTTCAGACTTTCAGAGAAACCCAGT 720  
Db 821 GAGCGCGAGATGCTGGCAATGACACAAAGAAACACGACTTCAGACTTTCAGAGAAACCCAGT 880  
Qy 721 GAAGGAGCTTAACCTACCAATGAATGCTGCTCATATTAAGAGAAACCAAGAGAGGTTATA 780  
Db 881 GAAGGAGCTTAACCTACCAATGAATGCTGCTCATATTAAGAGAAACCAAGAGAGGTTAAC 940  
Qy 781 TGTTTGGTCTCTAAATATCTTGGATTTGATATGAACCAACACATAGTCTCTTGTGTGCT 840  
Db 941 TGTTTGGTCTCT- -GATCTTGGATTTGATATGAACCAAC- - -AGTCTTGTGTGCTAC 994  
Qy 841 TGACAGAACCCAGTTTGTATGTACATATTCATATTCCTCTCTCTGTGTGTGTTTCGGGGGG 900  
Db 995 TGACAAAGCCCGAGTGTGTGTGCTATTT---ACATTCCTCTCTCTGTGTGTGTTTCGGGGGG 1050  
Qy 901 AAAAGACATTTTAGCCCTTTTAAAGTACTGATTAATTCATGTTATTTGGTGTGCTAT 960  
Db 1051 AAAAGAGACATTTTAAACCTTTTAAAGGTTATTTGACTTAATTTATGTTATTTGGTGTGCTAT 1110  
Qy 961 GAAAGTTGCCCTTAACCACTAAGGATTAATCAAGATTTTTCGCGAGACTTATACATGCTCTAG 1020  
Db 1111 GAAAGTTGCCCTTAACCACTAAGGATTAATCAAGATTTT---GCAGACTTCTCCGTGTCTAG 1167  
Qy 1021 GATC 1024  
Db 1168 GCTC 1171

RESULT 7  
US-10-152-661-541  
; Sequence 541, Application US/10152661  
; Publication No. US2003002835A1  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Murison, James G.  
; APPLICANT: Kumble, Krishanand D.  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; FILE REFERENCE: 11000.1011c5  
; CURRENT APPLICATION NUMBER: US/10/152,661  
; CURRENT FILING DATE: 2002-05-20  
; PRIOR APPLICATION NUMBER: 09/866,050  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 60/221,232  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: 60/206,650  
; PRIOR FILING DATE: 2000-05-24  
; PRIOR APPLICATION NUMBER: 09/312,283  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: PCT/NZ99/00051  
; PRIOR FILING DATE: 1999-04-29  
; PRIOR APPLICATION NUMBER: 09/188,930  
; PRIOR FILING DATE: 1998-11-09  
; PRIOR APPLICATION NUMBER: 09/069,726  
; PRIOR FILING DATE: 1998-04-29  
; NUMBER OF SEQ ID NOS: 725  
; SOFTWARE: FastSeq for Windows Version 4.0





5 1013 98.9 1830 50 US-10-342-887-1603 Sequence 1603, Ap  
6 1013 98.9 1998 26 US-09-606-776-5006 Sequence 5006, Ap  
7 1013 98.9 2342 28 US-09-644-869-3226 Sequence 9226, Ap  
8 1013 98.9 2342 28 US-09-644-870-9669 Sequence 9669, Ap  
9 1013 98.9 2342 28 US-09-649-162-9840 Sequence 9840, Ap  
10 1013 98.9 2342 28 US-09-652-109-9938 Sequence 9938, Ap  
11 1013 98.9 2342 28 US-09-652-816-9425 Sequence 9425, Ap  
12 1013 98.9 2342 30 US-09-698-013-6892 Sequence 6892, Ap  
13 1013 98.9 2342 32 US-09-726-175-3077 Sequence 3077, Ap  
14 1013 98.9 2342 32 US-09-726-806-5292 Sequence 5292, Ap  
15 1012 98.8 1909 33 US-09-770-173-2543 Sequence 2543, Ap  
16 1012 98.8 1973 27 US-09-629-469A-11306 Sequence 11306, A  
17 1012 98.8 2669 20 US-09-371-168-7090 Sequence 7090, Ap  
18 1012 98.8 2669 28 US-09-644-867-5952 Sequence 5952, Ap  
19 1012 98.8 2669 28 US-09-644-868-7420 Sequence 7420, Ap  
20 1012 98.8 2669 28 US-09-644-869-7588 Sequence 7588, Ap  
21 1012 98.8 2669 28 US-09-644-870-7347 Sequence 7347, Ap  
22 1012 98.8 2669 28 US-09-644-871-7477 Sequence 7477, Ap  
23 1012 98.8 2669 28 US-09-649-162-8302 Sequence 8302, Ap  
24 1012 98.8 2669 28 US-09-649-164-7849 Sequence 7849, Ap  
25 1012 98.8 2669 28 US-09-652-109-8037 Sequence 8037, Ap  
26 1012 98.8 2669 28 US-09-652-126-8504 Sequence 8504, Ap  
27 1012 98.8 2669 28 US-09-652-814-9037 Sequence 9037, Ap  
28 1012 98.8 2669 28 US-09-652-816-8405 Sequence 8405, Ap  
29 1012 98.8 2669 28 US-09-652-913-8643 Sequence 8643, Ap  
30 1012 98.8 2669 28 US-09-652-914-7662 Sequence 7662, Ap  
31 1012 98.8 2669 34 US-09-801-833-7090 Sequence 7090, Ap  
32 1012 98.8 3193 20 US-09-396-885-4251 Sequence 4251, Ap  
33 1012 98.8 3193 20 US-09-396-885A-4251 Sequence 4251, Ap  
34 1012 98.8 3193 20 US-09-399-932-4843 Sequence 4843, Ap  
35 1000 97.7 1898 1 PCT-US00-22350-22 Sequence 22, Appl  
36 1000 97.7 1898 2 PCT-US00-22350-22 Sequence 22, Appl  
37 1000 97.7 1898 33 US-09-768-826-22 Sequence 22, Appl  
38 985.6 96.2 1919 1 PCT-US00-05882-145 Sequence 145, Appl  
39 985.6 96.2 1919 2 PCT-US00-05882-145 Sequence 145, Appl  
40 985.6 96.2 1919 39 US-09-925-301-145 Sequence 145, Appl  
41 980 95.7 3805 15 US-09-925-301-145 Sequence 4, Appl  
42 980 95.7 3805 48 US-10-247-671-8 Sequence 8, Appl  
43 980 95.7 3805 87 US-60-323-784-8 Sequence 8, Appl  
44 980 95.7 4711 82 US-60-278-232-5507 Sequence 6507, Ap  
45 980 95.7 4711 87 US-60-324-185-21767 Sequence 21767, A

ALIGNMENTS

RESULT 1  
US-09-787-328A-3  
; Sequence 3, Application US/09787328A  
; GENERAL INFORMATION:  
; APPLICANT: Yu, Long  
; APPLICANT: Zhang, Honglai  
; APPLICANT: Fu, Qiang  
; APPLICANT: Zhao, Yong  
; APPLICANT: Tu, Qiang  
; TITLE OF INVENTION: NEW HUMAN HEPATOMA-DERIVED GROWTH FACTOR ENCODING SEQUENCE AND TH  
; FILE REFERENCE: 9548.50USWO  
; CURRENT APPLICATION NUMBER: US/09/787,328A  
; CURRENT FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: PCT/CN99/00139  
; PRIOR FILING DATE: 1999-09-06  
; PRIOR APPLICATION NUMBER: CN 98119758.2  
; PRIOR FILING DATE: 1998-09-22  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 1024  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-787-328A-3

Query Match 100.0%; Score 1024; DB 33; Length 1024;

Best Local Similarity 100.0%; Pred. No. 3.6e-223;  
Matches 1024; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACCGCTGTCGCGCGCGCTTGAAGCCGCGGGAGCGCGCAATTGTCGCGCGCGGG 60  
DB 1 ACCGCTGTCGCGCGCGCTTGAAGCCGCGGGAGCGCGCAATTGTCGCGCGCGGG 60  
QY 61 GGGGCGGCTCCCGGCACTTTCGCGGGACCAAGGACTTACCAGGAAGGGAGCGCTGG 120  
DB 61 GGGGCGGCTCCCGGCACTTTCGCGGGACCAAGGACTTACCAGGAAGGGAGCGCTGG 120  
QY 121 ATGCGCGCTCCCGCGCGCGAGTACAAAGCGGGCGACCTGGTCTTCGCCAAGATGA 180  
DB 121 ATGCGCGCTCCCGCGCGCGAGTACAAAGCGGGCGACCTGGTCTTCGCCAAGATGA 180  
QY 181 GGCTACCGCGACTCGGCGCGCGGATGTAATGAACTCCAGAGGGCGCTGTGAAGCTCA 240  
DB 181 GGCTACCGCGACTCGGCGCGCGGATGTAATGAACTCCAGAGGGCGCTGTGAAGCTCA 240  
QY 241 GCAACAGATATCTATCTCTTTTGGCACCCATGAATGCAATTTCTAGTCCCAAA 300  
DB 241 GCAACAGATATCTATCTCTTTTGGCACCCATGAATGCAATTTCTAGTCCCAAA 300  
QY 301 GACCTTTTCCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAGA 360  
DB 301 GACCTTTTCCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAGA 360  
QY 361 TTTAAAGAGGATTTGGGAAATAGAAATAACCCAGGAGTAAAGTTTACTGGCTACC 420  
DB 361 TTTAAAGAGGATTTGGGAAATAGAAATAACCCAGGAGTAAAGTTTACTGGCTACC 420  
QY 421 GCAATTCAGCAACAGAGCTCTTCAGAACTCAGGAGGAGGAGGAGGAGTCAATGCA 480  
DB 421 GCAATTCAGCAACAGAGCTCTTCAGAACTCAGGAGGAGGAGGAGGAGTCAATGCA 480  
QY 481 AGCAGTGAAGAGGAGTATAGAGTGAAGAGAGTGAAGAGGAGGAGGAGGAGGAGT 540  
DB 481 AGCAGTGAAGAGGAGTATAGAGTGAAGAGAGTGAAGAGGAGGAGGAGGAGTGA 540  
QY 541 AAAGAGGCTCAAAACGGAAGAGTATATCTTCAAGAAATCCTCTAAACAGTCCCG 600  
DB 541 AAAGAGGCTCAAAACGGAAGAGTATATCTTCAAGAAATCCTCTAAACAGTCCCG 600  
QY 601 AAATCTCCAGAGATGAAGTACAAAGAGTGAAGAGGAGGAGGAGGAGGAGGAGTCT 660  
DB 601 AAATCTCCAGAGATGAAGTACAAAGAGTGAAGAGGAGGAGGAGGAGGAGGAGTCT 660  
QY 661 GAGGCTGAGATGCGGCGCAACGACACAAAGAACCACTTCAGACTTTCGAGAAACCA 720  
DB 661 GAGGCTGAGATGCGGCGCAACGACACAAAGAACCACTTCAGACTTTCGAGAAACCA 720  
QY 721 GAAGGACCTTAACCTACCAATATGAATGCTGATATTAAGAGAAACCAAGAGGTTAT 780  
DB 721 GAAGGACCTTAACCTACCAATATGAATGCTGATATTAAGAGAAACCAAGAGGTTAT 780  
QY 781 TGTGTTGTTGCTAATATCTTGGATTTGATATCAACCAACACATAGTCTTGTGTCT 840  
DB 781 TGTGTTGTTGCTAATATCTTGGATTTGATATCAACCAACACATAGTCTTGTGTCT 840  
QY 841 TGACAGAACCCAGTTTGTATGATATTAATTCATATTCCTCTCTGTTGTTTGGGGGG 900  
DB 841 TGACAGAACCCAGTTTGTATGATATTAATTCATATTCCTCTCTGTTGTTTGGGGGG 900  
QY 901 AAAAGACATTTAGCTTTTAAAGTTACTGATTAATTTCAATTTGTTGTTGTTGAT 960  
DB 901 AAAAGACATTTAGCTTTTAAAGTTACTGATTAATTTCAATTTGTTGTTGTTGAT 960  
QY 961 GAAGTTGCCCTTAACCACTAAGGATATCAAGATTTTGGCAGACTTATACATGCTCT 1020  
DB 961 GAAGTTGCCCTTAACCACTAAGGATATCAAGATTTTGGCAGACTTATACATGCTCT 1020  
QY 1021 GATC 1024  
|||||

Db	1021	GATC 1024	
RESULT 2			
US-09-787-328B-3			
; Sequence 3, Application US/09787328B			
; GENERAL INFORMATION:			
; APPLICANT: Yu, Long			
; APPLICANT: Zhang, Honglai			
; APPLICANT: Fu, Qiang			
; APPLICANT: Zhao, Yong			
; APPLICANT: Tu, Qiang			
; TITLE OF INVENTION: NEW HUMAN HEPATOMA-DERIVED GROWTH FACTOR ENCODING SEQUENCE AND			
; TITLE OF INVENTION: POLYPEPTIDE ENCODED BY SUCH DNA SEQUENCE AND PRODUCING METHOD TH			
; FILE REFERENCE: 9548.50USWO			
; CURRENT APPLICATION NUMBER: US/09/787,328B			
; PRIOR FILING DATE: 2001-03-16			
; PRIOR APPLICATION NUMBER: PCT/CN99/00139			
; PRIOR FILING DATE: 1999-09-06			
; PRIOR APPLICATION NUMBER: CN 98119758.2			
; PRIOR FILING DATE: 1998-09-22			
; NUMBER OF SEQ ID NOS: 10			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 3			
; LENGTH: 1024			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-787-328B-3			
Query Match 100.0%; Score 1024; DB 33; Length 1024;			
Best Local Similarity 100.0%; Pred. No. 3.6e-223;			
Matches 1024; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ACCGCTCGTCCGCCGCTTGGAGCCGCGGGAGCGCGCAATTCGTCGCGCGCGG 60	
Db	1		
Qy	61	GGGGGGGCTCCCGGCATCTTCGGCGGACCAAGGACTACCAAGGAGGGAGCGGCTGG 120	
Db	61	GGGGGGGCTCCCGGCATCTTCGGCGGACCAAGGACTACCAAGGAGGGAGCGGCTGG 120	
Qy	121	ATGGCGGCTCGGGCCCGGAGTACAAAGCGGGCGACCTGGTCTTCGCCAAGATGAAG 180	
Db	121	ATGGCGGCTCGGGCCCGGAGTACAAAGCGGGCGACCTGGTCTTCGCCAAGATGAAG 180	
Qy	181	GGTACCCGCTCGGGCGGCGGATTTGATGATGATGATGATGATGATGATGATGATG 240	
Db	181	GGTACCCGCTCGGGCGGCGGATTTGATGATGATGATGATGATGATGATGATGATG 240	
Qy	241	GCAACAAGTATCTATCTTTTGGCACCCATGAACTGCATTTCTAGTCTCCAAA 300	
Db	241	GCAACAAGTATCTATCTTTTGGCACCCATGAACTGCATTTCTAGTCTCCAAA 300	
Qy	301	GACCTTTTCCATATAAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACCGAAGGA 360	
Db	301	GACCTTTTCCATATAAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACCGAAGGA 360	
Qy	361	TTTAAAGGAGGATTTGGGAAATAGAAATACCCAGGAGTAAAGTTTACTGCTACCAG 420	
Db	361	TTTAAAGGAGGATTTGGGAAATAGAAATACCCAGGAGTAAAGTTTACTGCTACCAG 420	
Qy	421	GCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAGGTTGGAATATCTGCAGATGCA 480	
Db	421	GCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAGGTTGGAATATCTGCAGATGCA 480	
Qy	481	AGCAGTACGAAAGAGGTGATAGATGAGTAAAGATGGAAGGCAAAAGAAAGATGAA 540	
Db	481	AGCAGTACGAAAGAGGTGATAGATGAGTAAAGATGGAAGGCAAAAGAAAGATGAA 540	
Qy	541	AAAGCAGGCTCAAAACGGAAGGATCATATCTTCAAGAAATCTCTTAAACAGTCCCGG 600	
Db	541	AAAGCAGGCTCAAAACGGAAGGATCATATCTTCAAGAAATCTCTTAAACAGTCCCGG 600	

Qy	601	AAATCTCCAGGAGATGAAGATGACAAAGACTTCGCAAGAGGAGGAAACCAAAAGCAGCTCT 660	
Db	601	AAATCTCCAGGAGATGAAGATGACAAAGACTTCGCAAGAGGAGGAAACCAAAAGCAGCTCT 660	
Qy	661	GAGGGTGGAGATGCGGGCAACGACACAAGAAACCAACTTCAGACTTTCGACAAAGGAGGTTATA 720	
Db	661	GAGGGTGGAGATGCGGGCAACGACACAAGAAACCAACTTCAGACTTTCGACAAAGGAGGTTATA 720	
Qy	721	GAGGGACCTAACTACCATTAATGCTGCTATATTAAGAGAAACCAACAAAGAGGTTATA 780	
Db	721	GAGGGACCTAACTACCATTAATGCTGCTATATTAAGAGAAACCAACAAAGAGGTTATA 780	
Qy	781	TGTTTGGTGTCTAATATCTTGGATTTGATATGACCAACACATAGTCTCTTGTGTCAT 840	
Db	781	TGTTTGGTGTCTAATATCTTGGATTTGATATGACCAACACATAGTCTCTTGTGTCAT 840	
Qy	841	TGACAGAACCCAGTTTGTATGATATTTTAAAGTTTACTGATTTAATTTTGGTTCAT 900	
Db	841	TGACAGAACCCAGTTTGTATGATATTTTAAAGTTTACTGATTTAATTTTGGTTCAT 900	
Qy	901	AAAAGACATTTTGGCTTTTAAAGTTTACTGATTTAATTTTGGTTCAT 960	
Db	901	AAAAGACATTTTGGCTTTTAAAGTTTACTGATTTAATTTTGGTTCAT 960	
Qy	961	GAAAGTTCCTTAAACCACTAAGGATTTTAAAGTTTTCGCGAGCTTATACATGCTAG 1020	
Db	961	GAAAGTTCCTTAAACCACTAAGGATTTTAAAGTTTTCGCGAGCTTATACATGCTAG 1020	
Qy	1021	GATC 1024	
Db	1021	GATC 1024	
RESULT 3			
PCT-US02-18947-1603			
; Sequence 1603, Application PC/TUS0218947			
; GENERAL INFORMATION:			
; APPLICANT: Rosetta Inpharmatics			
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients			
; FILE REFERENCE: 9301-175-228			
; CURRENT APPLICATION NUMBER: PCT/US02/18947			
; CURRENT FILING DATE: 2002-06-14			
; PRIOR APPLICATION NUMBER: 60/380,770			
; PRIOR FILING DATE: 2002-05-14			
; NUMBER OF SEQ ID NOS: 2699			
; SEQ ID NO 1603			
; LENGTH: 1830			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; PUBLICATION INFORMATION:			
; DATABASE ACCESSION NUMBER: NM 016073			
; DATABASE ENTRY DATE: 2001-06-18			
PCT-US02-18947-1603			
Query Match 98.9%; Score 1013; DB 1; Length 1830;			
Best Local Similarity 99.9%; Pred. No. 1.4e-220;			
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;			
Qy	1	ACCGCTCGTCCGCCGCTTGGAGCCCGGGAGCGCG-CCCAATTCGTCGCGCGCGG 59	
Db	4	ACCGCTCGTCCGCCGCTTGGAGCCCGGGAGCGCGCAATTCGTCGCGCGCGG 63	
Qy	60	GGGGGGGCGCTCCCGGCATCTTCGGCGGACCAAGGACTACCAAGGAGGGAGCGGCTGG 119	
Db	64	GGGGGGGCGCTCCCGGCATCTTCGGCGGACCAAGGACTACCAAGGAGGGAGCGGCTGG 123	
Qy	120	GATGCGGCGTCCCGGGCCCGCGAGTACAAAGCGGGCGACCTGGTCTTCGCCAAGATGAA 179	
Db	124	GATGCGGCGTCCCGGGCCCGCGAGTACAAAGCGGGCGACCTGGTCTTCGCCAAGATGAA 183	
Qy	180	GGGCTACCGCACTGGCGCGCGCGGATTTGATGAACTCCCGAGGCGGCTGTGAAGCTCC 239	
Db	184	GGGCTACCGCACTGGCGCGCGCGGATTTGATGAACTCCCGAGGCGGCTGTGAAGCTCC 243	

```
Qy 240 AGCAAAAGATATCTATCTCTCTTTTGGCACCCATGAAGTCAATCTAGTCCAA 299
|
|
|
Db 244 AGCAAAAGATATCTATCTCTTTTGGCACCCATGAAGTCAATCTAGTCCAA 303
|
|
|
Qy 300 AGACCTTTTCCATATAGAGTCAAGAGCAAGTTGGAAAGTCAACAAACGGAAGG 359
|
|
|
Db 304 AGACCTTTTCCATATAGAGTCAAGAGCAAGTTGGAAAGTCAACAAACGGAAGG 363
|
|
|
Qy 360 ATTTAAAGGATTTGGGAAATAGAAAATACCCAGGAGTAAAGTTTACGCTACCA 419
|
|
|
Db 364 ATTTAAAGGATTTGGGAAATAGAAAATACCCAGGAGTAAAGTTTACGCTACCA 423
|
|
|
Qy 420 GGCAATTCAGCAACAGAGCTCTTCAGAAATCGAGGAGAGGTTGAAATCTGCAGATGC 479
|
|
|
Db 424 GGCAATTCAGCAACAGAGCTCTTCAGAAATCGAGGAGAGGTTGAAATCTGCAGATGC 483
|
|
|
Qy 480 AAGCAGTGAGGAAGAGTATAGATGAGAGAGAGTGGAAAGGCAAAAGAAAGATGA 539
|
|
|
Db 484 AAGCAGTGAGGAAGAGTATAGATGAGAGAGAGTGGAAAGGCAAAAGAAAGATGA 543
|
|
|
Qy 540 AAAAGCAGGCTCAAAACGGAAGAGTATATATCTTCAAGAAATCCTTAAACAGTCCG 599
|
|
|
Db 544 AAAAGCAGGCTCAAAACGGAAGAGTATATATCTTCAAGAAATCCTTAAACAGTCCG 603
|
|
|
Qy 600 GAAATCTCCAGAGATGAAGATGACAAAGACTGCAAGAGAGAGGAAACAAAGCAGCTC 659
|
|
|
Db 604 GAAATCTCCAGAGATGAAGATGACAAAGACTGCAAGAGAGAGGAAACAAAGCAGCTC 663
|
|
|
Qy 719 TGAGGGTGAGATCGGGCAACGACACAAAGAAACAACTTCAGACTTGCAGAAACCCAG 723
|
|
|
Db 724 TGAGGGTGAGATCGGGCAACGACACAAAGAAACAACTTCAGACTTGCAGAAACCCAG 723
|
|
|
Qy 780 ATGTTTGGTCTTAATATCTCGATTGATATGAACCAACATAGTCTCTGTCTGTCGA 839
|
|
|
Db 784 ATGTTTGGTCTTAATATCTCGATTGATATGAACCAACATAGTCTCTGTCTGTCGA 843
|
|
|
Qy 840 TTGACAGAACCCAGTTGTATGATATATTCATATTCCTCTCTGTGTGTGTTTCGGGG 899
|
|
|
Db 844 TTGACAGAACCCAGTTGTATGATATATTCATATTCCTCTCTGTGTGTGTTTCGGGG 903
|
|
|
Qy 900 GAAAGACATTTAGCCTTTTAAAGTACTGATTAATTTTCAATGTTTGGTTGCA 959
|
|
|
Db 904 GAAAGACATTTAGCCTTTTAAAGTACTGATTAATTTTCAATGTTTGGTTGCA 963
|
|
|
Qy 960 TGAAGTTGCCCTTAACCACTAAGGATATCAAGATTTTGGCAGACTTATACATGCTA 1019
|
|
|
Db 964 TGAAGTTGCCCTTAACCACTAAGGATATCAAGATTTTGGCAGACTTATACATGCTA 1023
|
|
|
Qy 1020 GGATC 1024
|
|
|
Db 1024 GGATC 1028
|
|
|
```

## RESULT 4

```
US-10-172-118-1603
; Sequence 1603, Application US/10172118
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
```

```
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1603
; LENGTH: 1830
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_016073
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1603
```

```
Query Match 98.9%; Score 1013; DB 47; Length 1830;
Best Local Similarity 99.9%; Pred. No. 1.4e-220;
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
Qy 1 ACCGCTGTCGCCCGCTTGGAGCGCGCGGAGGCGG-CGCAATTCGTGCGCCCGCG 59
|
|
|
Db 4 ACCGCTGTCGCCCGCTTGGAGCGCGCGGAGGCGGCGCAATTCGTGCGCCCGCG 63
|
|
|
Qy 60 GGGGCGGCTCCCGGCTCTTCGCGGCGACCAAGNCTACCAAGGAGGAGCGGCTGG 119
|
|
|
Db 64 GGGGCGGCTCCCGGCTCTTCGCGGCGACCAAGNCTACCAAGGAGGAGGCGGCTGG 123
|
|
|
Qy 120 GATGGCGGCTCCCGGCGCGCGGAGTACAAAGCGGCGACCTGGTCTTCGCAAGATGAA 179
|
|
|
Db 124 GATGGCGGCTCCCGGCGCGCGGAGTACAAAGCGGCGACCTGGTCTTCGCAAGATGAA 183
|
|
|
Qy 180 GGGCTACCCGCTACGCGCGCGGATTTGATGAATCTCCAGAGGCGCTGTGAAGCTCC 239
|
|
|
Db 184 GGGCTACCCGCTACGCGCGCGGATTTGATGAATCTCCAGAGGCGCTGTGAAGCTCC 243
|
|
|
Qy 240 AGCAAAAGATATCTATCTCTTTTGGCACCCATGAAGTCAATCTAGTCCCA 299
|
|
|
Db 244 AGCAAAAGATATCTATCTCTTTTGGCACCCATGAAGTCAATCTAGTCCCA 303
|
|
|
Qy 300 AGACCTTTTCCATATAGAGTCAAGAGCAAGTTGGAAAGTCAACAAACGGAAGG 359
|
|
|
Db 304 AGACCTTTTCCATATAGAGTCAAGAGCAAGTTGGAAAGTCAACAAACGGAAGG 363
|
|
|
Qy 360 ATTTAAAGGATTTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACCA 419
|
|
|
Db 364 ATTTAAAGGATTTGGGAAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACCA 423
|
|
|
Qy 420 GCAATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTTGAAATCTGCAGATGC 479
|
|
|
Db 424 GCAATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTTGAAATCTGCAGATGC 483
|
|
|
Qy 480 AAGCAGTGAGGAAGAGTATAGAGTGAAGAGAGTGGAAAGGCAAAAGAAAGATGA 539
|
|
|
Db 484 AAGCAGTGAGGAAGAGTATAGAGTGAAGAGAGTGGAAAGGCAAAAGAAAGATGA 543
|
|
|
Qy 540 AAAAGCAGGCTCAAAACGGAAGAGTATATCTTCAAGAAATCCTTCAACAGTCCCG 599
|
|
|
Db 544 AAAAGCAGGCTCAAAACGGAAGAGTATATCTTCAAGAAATCCTTCAACAGTCCCG 603
|
|
|
Qy 600 GAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAGAGAGAGGTTGAAATCTGCAGATGC 659
|
|
|
Db 604 GAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAGAGAGAGGTTGAAATCTGCAGATGC 663
|
|
|
Qy 660 TGAGGGTGAGATCGGGCAACGACACAAAGAAACAACTTCAGACTTGCAGAAACCCAG 719
|
|
|
Db 664 TGAGGGTGAGATCGGGCAACGACACAAAGAAACAACTTCAGACTTGCAGAAACCCAG 723
|
|
|
Qy 720 TGAAGGAGCCTTAACCTACCAATGAATGCTGCTATTTAAGAGAAACCAAGAGGTTAT 779
|
|
|
Db 724 TGAAGGAGCCTTAACCTACCAATGAATGCTGCTATTTAAGAGAAACCAAGAGGTTAT 783
|
|
|
Qy 780 ATGTTTGGTCTTAATATCTTGGATTTGATATGAACCAACATAGTCTCTGTCTGTCGA 839
|
|
|
Db 784 ATGTTTGGTCTTAATATCTTGGATTTGATATGAACCAACATAGTCTCTGTCTGTCGA 843
|
|
|
```



QY 840 TTGACAGACCCAGGTTTGTATGATACATATTCATATTCCTCTCTGTGTGTTTCGGGG 899  
Db 844 TTGACAGACCCAGGTTTGTATGATACATATTCATATTCCTCTCTGTGTGTTTCGGGG 903  
QY 900 GAAAGACATTTAGCCTTTTAAAGTTTACTGATTTAATTTTCATGTTTGGTTGCA 959  
Db 904 GAAAGACATTTAGCCTTTTAAAGTTTACTGATTTAATTTTCATGTTTGGTTGCA 963  
QY 960 TGAAGTTGCCCTTAAACCACTAAGGATTTATCAAGATTTTTCGCGAGACTTATACATGCTCA 1019  
Db 964 TGAAGTTGCCCTTAAACCACTAAGGATTTATCAAGATTTTTCGCGAGACTTATACATGCTCA 1023  
QY 1020 GGATC 1024  
Db 1024 GGATC 1028

RESULT 5  
US-10-342-887-1603  
; Sequence 1603, Application US/10342887  
; GENERAL INFORMATION:  
; APPLICANT: Dai, Hongyue  
; APPLICANT: He, Yudong  
; APPLICANT: Linsley, Peter S.  
; APPLICANT: Mao, Mao  
; APPLICANT: Roberts, Christopher J.  
; APPLICANT: Van 't Veer, Laura Johanna  
; APPLICANT: Van de Vijver, Marc J.  
; APPLICANT: Bernards, Rene  
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
; FILE REFERENCE: 9301-188-999  
; CURRENT APPLICATION NUMBER: US/10/342,887  
; CURRENT FILING DATE: 2003-01-15  
; PRIOR APPLICATION NUMBER: 60/298,918  
; PRIOR FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: 60/380,710  
; PRIOR FILING DATE: 2002-05-14  
; PRIOR APPLICATION NUMBER: 10/172,118  
; PRIOR FILING DATE: 2002-06-14  
; NUMBER OF SEQ ID NOS: 2699  
; SEQ ID NO 1603  
; LENGTH: 1830  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-342-887-1603

Query Match 98.9%; Score 1013; DB 50; Length 1830;  
Best Local Similarity 99.9%; Pred. No. 1.4e-220;  
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ACCGCTCGTCCGCGCCGCTTGAGCCCGCGGGAGCGCG-CGCAATTTGTCGCGCCGCGG 59  
Db 4 ACCGCTCGTCCGCGCCGCTTGAGCCCGCGGGAGCGCGCGCAATTTGTCGCGCCGCGG 63  
QY 60 GGGGGCGGCTCCCGGCATCTTCGCGCGCACCAAGGATACCAAGGAGGGGAGCGGCTGG 119  
Db 64 GGGGGCGGCTCCCGGCATCTTCGCGCGCACCAAGGATACCAAGGAGGGGAGCGGCTGG 123  
QY 120 GATGGCGCTCCGCGCCCGCGGAGTACAAAGCGGGGACCTGCTTGGCCAAAGATGAA 179  
Db 124 GATGGCGCTCCGCGCCCGCGGAGTACAAAGCGGGGACCTGCTTGGCCAAAGATGAA 183  
QY 180 GGGCTACCCGCTCCGCGCCCGCGGATTTGATGAATCCCAAGGGGCGCTGTGAAGCTCC 239  
Db 184 GGGCTACCCGCTCCGCGCCCGCGGATTTGATGAATCCCAAGGGGCGCTGTGAAGCTCC 243  
QY 240 AGCAACAAGATCTCTATCTTTTGGCACCCTCATGAATCTTCTAGGTCCTCCAA 299  
Db 244 AGCAACAAGATCTCTATCTTTTGGCACCCTCATGAATCTTCTAGGTCCTCCAA 303  
QY 300 AGACCTTTTCCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACCGAAAG 359  
Db 304 AGACCTTTTCCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACCGAAAG 363

QY 360 ATTAAACGAAGGATTGTGGGAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACCA 419  
Db 364 ATTAAACGAAGGATTGTGGGAATAGAAAATAACCCAGGAGTAAAGTTTACTGGCTACCA 423  
QY 420 GGAATTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAAAGGTGGAAATCTTCAGATGC 479  
Db 424 GGAATTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAAAGGTGGAAATCTTCAGATGC 483  
QY 480 AGCAGTGGAGGAGAGGTGATAGTAGAAGAGATGAAAGGCAAGGCAAGAAAGAAATGA 539  
Db 484 AGCAGTGGAGGAGAGGTGATAGTAGAAGAGATGAAAGGCAAGGCAAGAAAGAAATGA 543  
QY 540 AAAAGCAGGCTCAAAAACGGAAAAAGTCAATATCTTCAAAAGAAATCTCTTAAACAGTCCCG 599  
Db 544 AAAAGCAGGCTCAAAAACGGAAAAAGTCAATATCTTCAAAAGAAATCTCTTAAACAGTCCCG 603  
QY 600 GAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAGAGAGAGAAAAACAAAAGCAGCTC 659  
Db 604 GAAATCTCCAGGAGATGAAGATGACAAAGACTGCAAGAGAGAGAAAAACAAAAGCAGCTC 663  
QY 660 TGAGGCTGAGATGCGGGCAACGACACAAAGAACACAACTTTCAGACTTTCAGAAAAACCCAG 719  
Db 664 TGAGGCTGAGATGCGGGCAACGACACAAAGAACACAACTTTCAGACTTTCAGAAAAACCCAG 723  
QY 720 TGAAGGACCTAACTACCATTAATGATGCTGCATATTTAAGAGAGAAACCAAGAAAGGTTAT 779  
Db 724 TGAAGGACCTAACTACCATTAATGATGCTGCATATTTAAGAGAGAAACCAAGAAAGGTTAT 783  
QY 780 ATGTTGGTTGTCTAATATCTTGGATTTGATATGAACCAACATAGTCTTGTGTTGTC 839  
Db 784 ATGTTGGTTGTCTAATATCTTGGATTTGATATGAACCAACATAGTCTTGTGTTGTC 843  
QY 840 TTGACAGACCCAGGTTTGTATGATACATATTCATATTCCTCTCTGTGTGTTTCGGGG 899  
Db 844 TTGACAGACCCAGGTTTGTATGATACATATTCATATTCCTCTCTGTGTGTTTCGGGG 903  
QY 900 GAAAGACATTTTAGCCTTTTAAAGTTTACTGATTTAATTTTCATGTTTGGTTGCA 959  
Db 904 GAAAGACATTTTAGCCTTTTAAAGTTTACTGATTTAATTTTCATGTTTGGTTGCA 963  
QY 960 TGAAGTTGCCCTTAAACCACTAAGGATTTATCAAGATTTTTCGCGAGACTTATACATGCTCA 1019  
Db 964 TGAAGTTGCCCTTAAACCACTAAGGATTTATCAAGATTTTTCGCGAGACTTATACATGCTCA 1023  
QY 1020 GGATC 1024  
Db 1024 GGATC 1028

RESULT 6  
US-09-606-776-5006  
; Sequence 5006, Application US/09606776  
; GENERAL INFORMATION:  
; APPLICANT: Holtzman, Douglas A.  
; APPLICANT: Myers, Paul  
; APPLICANT: Gearing, David P.  
; APPLICANT: Pan, Yang  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: 1600.1129-001  
; CURRENT APPLICATION NUMBER: US/09/606,776  
; PRIOR FILING DATE: 2000-06-27  
; PRIOR APPLICATION NUMBER: 60/141,578  
; PRIOR FILING DATE: 1999-06-29  
; PRIOR APPLICATION NUMBER: 60/141,379  
; PRIOR FILING DATE: 1999-06-28  
; PRIOR APPLICATION NUMBER: 60/141,138  
; PRIOR FILING DATE: 1999-06-28  
; PRIOR APPLICATION NUMBER: 60/141,581  
; PRIOR FILING DATE: 1999-06-29  
; NUMBER OF SEQ ID NOS: 5415  
; SOFTWARE: FastSeq for Windows Version 3.0

```
; SEQ ID NO 5006
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-606-776-5006

Query Match      98.9%; Score 1013; DB 26; Length 1998;
Best Local Similarity 99.9%; Pred. No. 1.5e-220;
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ACCGCTCGTCCGCGCGCTTGAGCCCGCGGGAGCGC-GCGCAATTCGTGGCCCGCGG 59
Db 13 ACCGCTCGTCCGCGCGCTTGAGCCCGCGGGAGCGCGGCGCAATTCGTGGCCCGCGG 72

QY 60 GGGGGGGGCTCCCGGCACTTCGCGGGGACCAAGGACTTACCAGAAAGGGGCGGCTGG 119
Db 73 GGGGGGGGCTCCCGGCACTTCGCGGGGACCAAGGACTTACCAGAAAGGGGCGGCTGG 132

QY 120 GATGGCGGTCCGCGCGCGCGGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 179
Db 133 GATGGCGGTCCGCGCGCGCGGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 192

QY 180 GGGCTACCCGCACTGGCGCGCGCGGATTTGATGAATCCCGAGGGCGCTGTGAAGCTCC 239
Db 193 GGGCTACCCGCACTGGCGCGCGCGGATTTGATGAATCCCGAGGGCGCTGTGAAGCTCC 252

QY 240 AGCAAAACAGTATCCTATCTCTTTTGGGACCCCATGAACTGCAATCTTAGGTCCCAA 299
Db 253 AGCAAAACAGTATCCTATCTCTTTTGGGACCCCATGAACTGCAATCTTAGGTCCCAA 312

QY 300 AGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACAAACGAAAGG 359
Db 313 AGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACAAACGAAAGG 372

QY 360 ATTTAACGAGGATTTGCGGAATAGAAAATACCCAGGAGTAAAGTTTACTGGCTACCA 419
Db 373 ATTTAACGAGGATTTGCGGAATAGAAAATACCCAGGAGTAAAGTTTACTGGCTACCA 432

QY 420 GGCATTCAGCAACAGAGCTCTTCAGAACTCAGGAGAGGTGGAATATCTGCGATGTC 479
Db 433 GGCATTCAGCAACAGAGCTCTTCAGAACTCAGGAGAGGTGGAATATCTGCGATGTC 492

QY 480 AAGCAGTGAGGAAGAGGTGATAGAGTACAAAGAGATGAAAGGCAAAAGAAAGATGA 539
Db 493 AAGCAGTGAGGAAGAGGTGATAGAGTACAAAGAGATGAAAGGCAAAAGAAAGATGA 552

QY 540 AAAAGCAGCTCAAAACGGAAGAGTATATATCTTTCAAAGAAATCCTCTAAACAGTCCG 599
Db 553 AAAAGCAGCTCAAAACGGAAGAGTATATATCTTTCAAAGAAATCCTCTAAACAGTCCG 612

QY 600 GAAATCTCAGGAGATGAAGTACAAAGACTGCAAGAGAGGAAACAAAGCAGCTC 659
Db 613 GAAATCTCAGGAGATGAAGTACAAAGACTGCAAGAGAGGAAACAAAGCAGCTC 672

QY 660 TGAGGCTGGAGATGCGGGCAACGACACAAAGAAACAACTTCAGACTTTCAGAAACCCAG 719
Db 673 TGAGGCTGGAGATGCGGGCAACGACACAAAGAAACAACTTCAGACTTTCAGAAACCCAG 732

QY 720 TGAAGGGACCTTAACCTACCATATGAATGCTGATATTAAGAAACCAACAGAAAGTTAT 779
Db 733 TGAAGGGACCTTAACCTACCATATGAATGCTGATATTAAGAAACCAACAGAAAGTTAT 792

QY 780 ATGTTTGGTGTCTAATATCTTGAATTTGATATGAACCAACATAGTCTTGTGTGCA 839
Db 793 ATGTTTGGTGTCTAATATCTTGAATTTGATATGAACCAACATAGTCTTGTGTGCA 852

QY 840 TTGACAGAACCCAGTTTGTATGTATCATTTATTCATCTCTCTGTTGTGTTTCGGGG 899
Db 853 TTGACAGAACCCAGTTTGTATGTATCATTTATTCATCTCTCTGTTGTGTTTCGGGG 912

QY 900 GAAAGACATTTTAGCCTTTTAAAGAGTACTGATTTAATTTCAATGTTATTTGGTTGCA 959
Db 913 GAAAGACATTTTAGCCTTTTAAAGAGTACTGATTTAATTTCAATGTTATTTGGTTGCA 972
```

```
QY 960 TGAAGTTGCCCTTAACCACTAAGGATTATCAAGATTTTGGCAGACTTATACATGTCTA 1019
Db 973 TGAAGTTGCCCTTAACCACTAAGGATTATCAAGATTTTGGCAGACTTATACATGTCTA 1032

QY 1020 GGATC 1024
Db 1033 GGATC 1037
```

```
RESULT 7
US-09-644-869-9226
; Sequence 9226, Application US/09644869
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Monahan, John
; APPLICANT: Richardson, Jennifer
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.1182-001
; CURRENT APPLICATION NUMBER: US/09/644,869
; PRIOR FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: 60/151,062
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 9708
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9226
; LENGTH: 2342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-644-869-9226
```

```
Query Match      98.9%; Score 1013; DB 28; Length 2342;
Best Local Similarity 99.9%; Pred. No. 1.6e-220;
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
QY 1 ACCGCTCGTCCGCGCGCTTGAGCCCGCGGGAGCGC-GCGCAATTCGTGGCCCGCGG 59
Db 13 ACCGCTCGTCCGCGCGCTTGAGCCCGCGGGAGCGCGGCGCAATTCGTGGCCCGCGG 72

QY 60 GGGGGGGGCTCCCGGCACTTCGCGGGGACCAAGGACTTACAGGAAGGGGAGCGGCTGG 119
Db 73 GGGGGGGGCTCCCGGCACTTCGCGGGGACCAAGGACTTACAGGAAGGGGAGCGGCTGG 132

QY 120 GATGGCGGTCCGCGCGCGCGGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 179
Db 133 GATGGCGGTCCGCGCGCGCGGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 192

QY 180 GGGCTACCCGCACTGGCGCGCGCGGATTTGATGAATCCCGAGGGCGCTGTGAAGCTCC 239
Db 193 GGGCTACCCGCACTGGCGCGCGCGGATTTGATGAATCCCGAGGGCGCTGTGAAGCTCC 252

QY 240 AGCAAAACAGTATCCTATCTCTTTTGGGACCCCATGAAACTGCAATTCCTAGGTCCCAA 299
Db 253 AGCAAAACAGTATCCTATCTCTTTTGGGACCCCATGAAACTGCAATTCCTAGGTCCCAA 312

QY 300 AGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACAAACGAAAGG 359
Db 313 AGACCTTTTCCATATAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACAAACGAAAGG 372

QY 360 ATTTAACGAGGATTTGGGAATAGAAAATACCCAGGAGTAAAGTTTACTGGCTACCA 419
Db 373 ATTTAACGAGGATTTGGGAATAGAAAATACCCAGGAGTAAAGTTTACTGGCTACCA 432

QY 420 GGCATTCAGCAACAGAGCTCTTCAGAACTCAGGAGAGGTGGAATATCTGCGATGTC 479
Db 433 GGCATTCAGCAACAGAGCTCTTCAGAACTCAGGAGAGGTGGAATATCTGCGATGTC 492

QY 480 AAGCAGTGAGGAAGAGGTGATAGAGTACAAAGAGATGAAAGGCAAAAGAAAGATGA 539
Db 493 AAGCAGTGAGGAAGAGGTGATAGAGTACAAAGAGATGAAAGGCAAAAGAAAGATGA 552
```

Qy 540 AAAAGCAGGCTCAAAAAGGAAAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCG 599  
Db 553 AAAAGCAGGCTCAAAAAGGAAAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCG 612  
Qy 600 GAAATCTCCAGGAGATGAGATGACAAAGATGCAAGAGAGAGAGAGAGAGAGAGAGAG 659  
Db 613 GAAATCTCCAGGAGATGAGATGACAAAGATGCAAGAGAGAGAGAGAGAGAGAGAGAG 672  
Qy 660 TGAGGCTGGAGATGCGGCAACGACACAAAGAAACCAACTTTCAGACTTTCAGAGAAAC 719  
Db 673 TGAGGCTGGAGATGCGGCAACGACACAAAGAAACCAACTTTCAGACTTTCAGAGAAAC 732  
Qy 720 TGAAGGACCTTAACCTAATGAATGCTGCTCATATTAAGAGAGAGAGAGAGAGAGAGAG 779  
Db 733 TGAAGGACCTTAACCTAATGAATGCTGCTCATATTAAGAGAGAGAGAGAGAGAGAGAG 792  
Qy 780 ATGTTGGTGTCTTAATATCTTGGATTTGATATGAACCAACACATAGTCTTGTGTCA 839  
Db 793 ATGTTGGTGTCTTAATATCTTGGATTTGATATGAACCAACACATAGTCTTGTGTCA 852  
Qy 840 TTGACAGAACCCAGTTTGTATGATATTAATTTTCAATTTTCAATTTTCAATTTTCAAT 899  
Db 853 TTGACAGAACCCAGTTTGTATGATATTAATTTTCAATTTTCAATTTTCAATTTTCAAT 912  
Qy 900 GAAAAGACATTTTGGCTTTTAAAAAGTTTAACTGATTTTAACTGATTTTAACTGATTT 959  
Db 913 GAAAAGACATTTTGGCTTTTAAAAAGTTTAACTGATTTTAACTGATTTTAACTGATTT 972  
Qy 960 TGAAGTTGCCCTTAACCACTAAGGATTTTAAAGATTTTAAAGATTTTAAAGATTTTAA 1019  
Db 973 TGAAGTTGCCCTTAACCACTAAGGATTTTAAAGATTTTAAAGATTTTAAAGATTTTAA 1032  
Qy 1020 GGATC 1024  
Db 1033 GGATC 1037

## RESULT 8

US-09-644-870-9669

; Sequence 9669, Application US/09644870

; GENERAL INFORMATION:

; APPLICANT: DiStefano, Peter

; APPLICANT: Holtzman, Douglas A.

; APPLICANT: Curtis, Rory A.J.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES

; TITLE OF INVENTION: THEREFOR

; FILE REFERENCE: 1600.1172-001

; CURRENT APPLICATION NUMBER: US/09/644,870

; CURRENT FILING DATE: 2000-08-28

; PRIOR APPLICATION NUMBER: 60/150,752

; PRIOR FILING DATE: 1999-08-26

; NUMBER OF SEQ ID NOS: 10434

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 9669

; LENGTH: 2342

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-644-870-9669

Query Match 98.9%; Score 1013; DB 28; Length 2342;  
Best Local Similarity 99.9%; Pred. No. 1.6e-220;  
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
Qy 1 ACCGCTGTCCGCGCGGCTTGAAGCGCGCGGAGAGCGG-GCGCAATTCGTGCGCGCGCGG 59  
Db 13 ACCGCTGTCCGCGCGGCTTGAAGCGCGCGGAGAGCGGCGCAATTCGTGCGCGCGCGG 72  
Qy 60 GGGGGCGGCTTCCGCGGATCTTCGCGGACCAAGGACTACCAAGAGGGAGCGGCTGG 119  
Db 73 GGGGGCGGCTTCCGCGGATCTTCGCGGACCAAGGACTACCAAGAGGGAGCGGCTGG 132  
Qy 120 GATGGCGGTCCGCGCGCGCGCGAGTACAAAGCGGCGGACCTGTGCTTTCGCCAAGATGAA 179

Db 133 GATGGCGGTCCGCGCGCGCGAGTACAAAGCGGCGGACCTGTGCTTTCGCCAAGATGAA 192  
Qy 180 GGGCTACCCGCACTTGGCGCGCGCGGATTTGATGAATCCAGAGAGCGGCTGTGAAGCGCTCC 239  
Db 193 GGGCTACCCGCACTTGGCGCGCGCGGATTTGATGAATCCAGAGAGCGGCTGTGAAGCGCTCC 252  
Qy 240 AGCAAAACAAGTATCTTCTTTTGGCAGCCCATGAAACTGCAATTTCTAGGTCCCAA 299  
Db 253 AGCAAAACAAGTATCTTCTTTTGGCAGCCCATGCAATTTCTAGGTCCCAA 312  
Qy 300 AGACCTTTTTCATATAGGAGTACAAAGACAAGTTTGGAAAGTTCAAACAAACCGAAAGG 359  
Db 313 AGACCTTTTTCATATAGGAGTACAAAGACAAGTTTGGAAAGTTCAAACAAACCGAAAGG 372  
Qy 360 ATTTAAACGAAGGATTTGGGAAATAGAAAATACCCAGAGAGTAAAGTTTACTGGCTACCA 419  
Db 373 ATTTAAACGAAGGATTTGGGAAATAGAAAATACCCAGAGAGTAAAGTTTACTGGCTACCA 432  
Qy 420 GGCAATTCAGCAACAGAGCTCTTTCAGAAACTGAGGGAGAGAGTGGAAATATCTGCAGATGC 479  
Db 433 GGCAATTCAGCAACAGAGCTCTTTCAGAAACTGAGGGAGAGAGTGGAAATATCTGCAGATGC 492  
Qy 480 AAGCAGTGAAGAGAGAGTGTATAGATAGAGAGAGATGGAAAGGCGCAAAAGAAAGATGA 539  
Db 493 AAGCAGTGAAGAGAGAGTGTATAGATAGAGAGAGATGGAAAGGCGCAAAAGAAAGATGA 552  
Qy 540 AAAAGCAGGCTCAAAACGGGAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCCG 599  
Db 553 AAAAGCAGGCTCAAAACGGGAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCCG 612  
Qy 600 GAAATCTCCAGGAGATGCAAGATGCAAGAGTGCAGAAAGAGAGAGAGAGAGAGAGAGAG 659  
Db 613 GAAATCTCCAGGAGATGCAAGATGCAAGAGTGCAGAAAGAGAGAGAGAGAGAGAGAGAG 672  
Qy 660 TGAGGCTGGAGATGCGGCGCAACGACACAAAGAAACCAACTTTCAGACTTTCAGAGAAAC 719  
Db 673 TGAGGCTGGAGATGCGGCGCAACGACACAAAGAAACCAACTTTCAGACTTTCAGAGAAAC 732  
Qy 720 TGAAGGACCTTAACCTAATGAATGCTGCTCATATTAAGAGAGAGAGAGAGAGAGAGAG 779  
Db 733 TGAAGGACCTTAACCTAATGAATGCTGCTCATATTAAGAGAGAGAGAGAGAGAGAGAG 792  
Qy 780 ATGTTGGTGTCTTAATATCTTGGATTTGATATGAACCAACACATAGTCTTGTGTCA 839  
Db 793 ATGTTGGTGTCTTAATATCTTGGATTTGATATGAACCAACACATAGTCTTGTGTCA 852  
Qy 840 TTGACAGAACCCAGTTTGTATGATATTAATTTTCAATTTTCAATTTTCAATTTTCAAT 899  
Db 853 TTGACAGAACCCAGTTTGTATGATATTAATTTTCAATTTTCAATTTTCAATTTTCAAT 912  
Qy 900 GAAAAGACATTTTAGCCTTTTAAAAAGTTTAAAGTTCATTTTAAAGTTCATTTTAAAG 959  
Db 913 GAAAAGACATTTTAGCCTTTTAAAAAGTTTAAAGTTCATTTTAAAGTTCATTTTAAAG 972  
Qy 960 TGAAGTTGCCCTTAACCACTAAGGATTTTAAAGATTTTAAAGATTTTAAAGATTTTAA 1019  
Db 973 TGAAGTTGCCCTTAACCACTAAGGATTTTAAAGATTTTAAAGATTTTAAAGATTTTAA 1032  
Qy 1020 GGATC 1024  
Db 1033 GGATC 1037

## RESULT 9

US-09-649-162-9840

; Sequence 9840, Application US/09649162

; GENERAL INFORMATION:

; APPLICANT: Shyjan, Andrew W.

; APPLICANT: Richardson, Jennifer

; APPLICANT: Holtzman, Douglas A.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES

; TITLE OF INVENTION: THEREFOR

```
; FILE REFERENCE: 1600.1181-001
; CURRENT APPLICATION NUMBER: US/09/649,162
; CURRENT FILING DATE: 2000-08-28
; PRIOR APPLICATION NUMBER: 60/151,057
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 9990
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9840
; LENGTH: 2342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-649-162-9840

Query Match      98.9%; Score 1013; DB 28; Length 2342;
Best Local Similarity 99.9%; Pred. No. 1.6e-220;
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ACCGCTCGTCCGCCCGCTTACGCGCGCGGAGCGC-GCGCAATTCGTGCGCGCGCGG 59
Db 13 ACCGCTCGTCCGCCCGCTTACGCGCGCGGAGCGCGCAATTCGTGCGCGCGCGG 72
QY 60 GGGGGCGGCTCCCGGCATCTTCGCGCGACCAAGACTACCAGGAGGGAGCGGCTGG 119
Db 73 GGGGGCGGCTCCCGGCATCTTCGCGCGACCAAGACTACCAGGAGGGAGCGGCTGG 132
QY 120 GATGGCGGCTCCCGCGCGCGCGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 179
Db 133 GATGGCGGCTCCCGCGCGCGCGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 192
QY 180 GGGCTACCCGCACTCGCGCGCGCGCGGATGATGAACTCCCGAGGGCGCTGTGAAGCTCC 239
Db 193 GGGCTACCCGCACTCGCGCGCGCGCGGATGATGAACTCCCGAGGGCGCTGTGAAGCTCC 252
QY 240 AGCAATTCAGCACAGAGCTCTTCAGAACTGAGGAGAGGTGGAATTAAGTCTGCAGATGC 479
Db 433 GGCATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTGGAATTAAGTCTGCAGATGC 492
QY 480 AAGCAGTGGAGGAAGAGGTGATAGAGTACAAAGAGATGAAAGCGCAAGAAAGAAATGA 539
Db 493 AAGCAGTGGAGGAAGAGGTGATAGAGTACAAAGAGATGAAAGCGCAAGAAAGAAATGA 552
QY 540 AAAAGCAGGCTCAAAACGAAAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCCG 599
Db 553 AAAAGCAGGCTCAAAACGAAAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCCG 612
QY 600 GAAATCTCCAGGAGTCAAGATGACAAAGACTGCAAGAGAGGAAAAACAAAGCAGCTC 659
Db 613 GAAATCTCCAGGAGTCAAGATGACAAAGACTGCAAGAGAGGAAAAACAAAGCAGCTC 672
QY 660 TGAGGTGGAGATGCGGGCAACGACACAAAGAACCAACTTCAGACTTCGAGAAACACAG 719
Db 673 TGAGGTGGAGATGCGGGCAACGACACAAAGAACCAACTTCAGACTTCGAGAAACACAG 732
QY 720 TGAAGGGAACCTAATACCAATATGAATGCTGCATATTAAGAGAAACCAACAAGGTTAT 779
Db 733 TGAAGGGAACCTAATACCAATATGAATGCTGCATATTAAGAGAAACCAACAAGGTTAT 792
QY 780 ATGTTTGGTGTCTAATATCTTGGATTTGATATGAAACCAACATAGTCTTGTGTCA 839
Db 793 ATGTTTGGTGTCTAATATCTTGGATTTGATATGAAACCAACATAGTCTTGTGTCA 852

RESULT 10
US-09-652-109-9938
; Sequence 9938, Application US/09652109
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1180-001
; CURRENT APPLICATION NUMBER: US/09/652,109
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,128
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 10105
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9938
; LENGTH: 2342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-652-109-9938

Query Match      98.9%; Score 1013; DB 28; Length 2342;
Best Local Similarity 99.9%; Pred. No. 1.6e-220;
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 ACCGCTCGTCCGCCCGCTTACGCGCGCGGAGCGC-GCGCAATTCGTGCGCGCGCGG 59
Db 13 ACCGCTCGTCCGCCCGCTTACGCGCGCGGAGCGCGCAATTCGTGCGCGCGCGG 72
QY 60 GGGGGCGGCTCCCGGCATCTTCGCGCGACCAAGACTACCAGGAGGGAGCGGCTGG 119
Db 73 GGGGGCGGCTCCCGGCATCTTCGCGCGACCAAGACTACCAGGAGGGAGCGGCTGG 132
QY 120 GATGGCGGCTCCCGCGCGCGCGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 179
Db 133 GATGGCGGCTCCCGCGCGCGCGAGTACAAAGCGGCGACCTGGTCTTCGCCAAGATGAA 192
QY 180 GGGCTACCCGCACTCGCGCGCGCGCGGATGATGAACTCCCGAGGGCGCTGTGAAGCTCC 239
Db 193 GGGCTACCCGCACTCGCGCGCGCGCGGATGATGAACTCCCGAGGGCGCTGTGAAGCTCC 252
QY 240 AGCAATTCAGCACAGAGCTCTTCAGAACTGAGGAGAGGTGGAATTAAGTCTGCAGATGC 479
Db 433 GGCATTCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTGGAATTAAGTCTGCAGATGC 492
```



; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES

; FILE REFERENCE: 1600.2013-001

; CURRENT APPLICATION NUMBER: US/09/698,013

; CURRENT FILING DATE: 2000-10-27

; PRIOR APPLICATION NUMBER: 60/162,360

; PRIOR FILING DATE: 1999-10-29

; NUMBER OF SEQ ID NOS: 7935

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6892

; LENGTH: 2342

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-698-013-6892

Query Match 98.9%; Score 1013; DB 30; Length 2342;

Best Local Similarity 99.9%; Pred. No. 1.6e-220;

Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
QY 1 ACCGCTCGTCCGCCCGGCTTGAGGCCCGCGGGAGCGC-GCGCAATTCTGTCGCCCGCGG 59
Db 13 ACCGCTCGTCCGCCCGGCTTGAGGCCCGCGGGAGCGGCGCAATTCTGTCGCCCGCGG 72
QY 60 GGGGGCGGCTCCCGGCACTCTTCGCGGCGACCAAGGACTTACAGGAAGGGAGCGGCTGG 119
Db 73 GGGGGCGGCTCCCGGCACTCTTCGCGGCGACCAAGGACTTACAGGAAGGGAGCGGCTGG 132
QY 120 GATGCGCGCTCCCGGCGGCGGCTTGAGGCCCGCGGGAGCGC-GCGCAATTCTGTCGCCCGGCAATGAA 179
Db 133 GATGCGCGCTCCCGGCGGCGGCTTGAGGCCCGCGGGAGCGGCGCAATTCTGTCGCCCGGCAATGAA 192
QY 180 GGGCTACCCGCACTGGCGGCGGCGGCTTGAGGAATGAACTCCAGAGGCGCTGTGAAGCTCC 239
Db 193 GGGCTACCCGCACTGGCGGCGGCGGCTTGAGGAATGAACTCCAGAGGCGCTGTGAAGCTCC 252
QY 240 AGCAACAAGTATCCTATCTCTTTTGGCACCCTGAACTGAACTGCTTCTAGGTCCCAA 299
Db 253 AGCAACAAGTATCCTATCTCTTTTGGCACCCTGAACTGAACTGCTTCTAGGTCCCAA 312
QY 300 AGACCTTTTCCATATAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACGAAAGG 359
Db 313 AGACCTTTTCCATATAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACGAAAGG 372
QY 360 ATTTAAAGAGGATTTGGGAAATAGAAATACCCAGGATGAAAGTTTACTGGCTACCA 419
Db 373 ATTTAAAGAGGATTTGGGAAATAGAAATACCCAGGATGAAAGTTTACTGGCTACCA 432
QY 420 GGCATTTCCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTTGAAATACTGCAAGTGC 479
Db 433 GGCATTTCCAGCAACAGAGCTCTTCAGAACTGAGGAGAGGTTGAAATACTGCAAGTGC 492
QY 480 AAGCAGTGAAGAGGATGATAGATAGAGTGAAGAGATGAGAAAGCGAAAGAAAGTGA 539
Db 493 AAGCAGTGAAGAGGATGATAGATAGAGTGAAGAGATGAGAAAGCGAAAGAAAGTGA 552
QY 540 AAAAGCAGCTCAAAACGAAAGTCTATATCTTCAAGAACTCAAGAACTCTTAAACGTCCG 599
Db 553 AAAAGCAGCTCAAAACGAAAGTCTATATCTTCAAGAACTCAAGAACTCTTAAACGTCCG 612
QY 600 GAAATCTCCAGGAGATGAGATGACAAAGACTGCAAGAGAGGAGGAAACAAAGCAGCTC 659
Db 613 GAAATCTCCAGGAGATGAGATGACAAAGACTGCAAGAGAGGAGGAAACAAAGCAGCTC 672
QY 660 TGAGGTGAGATGCGGGCAACGACACAAAGAAACCAACTTCAGACTTGCAGAAACGAG 719
Db 673 TGAGGTGAGATGCGGGCAACGACACAAAGAAACCAACTTCAGACTTGCAGAAACGAG 732
QY 720 TGAAGGACCTAACTACCATATGATGCTGATATTAAGGAAACCAACGAAAGTTAT 779
Db 733 TGAAGGACCTAACTACCATATGATGCTGATATTAAGGAAACCAACGAAAGTTAT 792
QY 780 ATGTTTGGTGTCTAATATCTTGGATTTGATGAAACCAACATAGTCTTGTGTCA 839
```

```
Db 793 ATGTTTGGTGTCTAATATCTTGGATTTGATGAAACCAACACATAGTCTTGTGTCA 852
QY 840 TTGACAGAACCCAGTTTGTATGATACATTTATTTATTTCTCTCTGTTGTGTTTCGGGG 899
Db 853 TTGACAGAACCCAGTTTGTATGATACATTTATTTATTTCTCTCTGTTGTGTTTCGGGG 912
QY 900 GAAAAGACATTTTAGCCCTTTTAAAAAGTTACTGATTTTAAATTTCAATTTATTTGGTTGCA 959
Db 913 GAAAAGACATTTTAGCCCTTTTAAAAAGTTACTGATTTTAAATTTCAATTTATTTGGTTGCA 972
QY 960 TGAAGTGGCTTTAACCACTAAGGATTTATCAAGATTTTGGCAGACTTATACATGCTA 1019
Db 973 TGAAGTGGCTTTAACCACTAAGGATTTATCAAGATTTTGGCAGACTTATACATGCTA 1032
QY 1020 GGATC 1024
Db 1033 GGATC 1037

RESULT 13
US-09-726-175-3077
; Sequence 3077, Application US/09726175
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Fraser, Christopher C.
; APPLICANT: Donovan, Michael J.
; APPLICANT: Holtzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2054-001
; CURRENT APPLICATION NUMBER: US/09/726,175
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/167,859
; PRIOR FILING DATE: 1999-11-29
; NUMBER OF SEQ ID NOS: 3770
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3077
; LENGTH: 2342
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-726-175-3077

Query Match 98.9%; Score 1013; DB 32; Length 2342;
Best Local Similarity 99.9%; Pred. No. 1.6e-220;
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 ACCGCTCGTCCGCCCGGCTTGAGGCCCGCGGGAGCGC-GCGCAATTCTGTCGCCCGCGG 59
Db 13 ACCGCTCGTCCGCCCGGCTTGAGGCCCGCGGGAGCGGCGCAATTCTGTCGCCCGCGG 72
QY 60 GGGGGCGGCTCCCGGCACTCTTCGCGGCGACCAAGGACTTACAGGAAGGGAGCGGCTGG 119
Db 73 GGGGGCGGCTCCCGGCACTCTTCGCGGCGACCAAGGACTTACAGGAAGGGAGCGGCTGG 132
QY 120 GATGCGCGCTCCCGGCGGCGGCTTGAGGCCCGCGGGAGCGC-GCGCAATTCTGTCGCCCGGCAATGAA 179
Db 133 GATGCGCGCTCCCGGCGGCGGCTTGAGGCCCGCGGGAGCGGCGCAATTCTGTCGCCCGGCAATGAA 192
QY 180 GGGCTACCCGCACTGGCGGCGGCGGCTTGAGGAATGAACTCCAGAGGCGCTGTGAAGCTCC 239
Db 193 GGGCTACCCGCACTGGCGGCGGCGGCTTGAGGAATGAACTCCAGAGGCGCTGTGAAGCTCC 252
QY 240 AGCAACAAGTATCCTATCTCTTTTGGCACCCTGAACTGAACTGCTTCTAGGTCCCAA 299
Db 253 AGCAACAAGTATCCTATCTCTTTTGGCACCCTGAACTGAACTGCTTCTAGGTCCCAA 312
QY 300 AGACCTTTTCCATATAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACGAAAGG 359
Db 313 AGACCTTTTCCATATAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACGAAAGG 372
QY 360 ATTTAAAGAGGATTTGGGAAATAGAAATACCCAGGATGAAAGTTTACTGGCTACCA 419
Db 373 ATTTAAAGAGGATTTGGGAAATAGAAATACCCAGGATGAAAGTTTACTGGCTACCA 432
```

QY 420 GGCAATTTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAAAGGTGGAAATACTGCAGATGC 479  
| | | | |  
Db 433 GGCAATTTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAAAGGTGGAAATACTGCAGATGC 492  
| | | | |  
QY 480 AAGCAGTGAAGAGAGGTGATAGAGTAGAGAGATGGAAGAGGCAAAAGAAAGATGA 539  
| | | | |  
Db 493 AAGCAGTGAAGAGAGGTGATAGAGTAGAGAGATGGAAGAGGCAAAAGAAAGATGA 552  
| | | | |  
QY 540 AAAAGCAGGCTCAAAACCGAAAAGTCAATATCTTCAAGAAATCTCTTAAACAGATCCCG 599  
| | | | |  
Db 553 AAAAGCAGGCTCAAAACCGAAAAGTCAATATCTTCAAGAAATCTCTTAAACAGATCCCG 612  
| | | | |  
QY 600 GAAATCTCCAGAGATGAAGATGACAAAGACTGCAAAAGAGAGAGAAACAAAAGCAGCTC 659  
| | | | |  
Db 613 GAAATCTCCAGAGATGAAGATGACAAAGACTGCAAAAGAGAGAGAAACAAAAGCAGCTC 672  
| | | | |  
QY 660 TGAGGTGGAGATGCGGGCAACGACACAGAAACAACTTCAGACTTCGAGAAAACCCAG 719  
| | | | |  
Db 673 TGAGGTGGAGATGCGGGCAACGACACAGAAACAACTTCAGACTTCGAGAAAACCCAG 732  
| | | | |  
QY 720 TGAAGGACCTAACTACCAATTAATGCTGCTGATATTAAGAGAAACCAAGAGAGTTAT 779  
| | | | |  
Db 733 TGAAGGACCTAACTACCAATTAATGCTGCTGATATTAAGAGAAACCAAGAGAGTTAT 792  
| | | | |  
QY 780 ATGTTGGTGTCTAATATCTTGGATTTGATATGAACCAACACATAGTCTCTTGTGTCA 839  
| | | | |  
Db 793 ATGTTGGTGTCTAATATCTTGGATTTGATATGAACCAACACATAGTCTCTTGTGTCA 852  
| | | | |  
QY 840 TTGACAGAACCCAGATTTGATGATATTAATTCCTCTCTGTTGTTGGGG 899  
| | | | |  
Db 853 TTGACAGAACCCAGATTTGATGATATTAATTCCTCTCTGTTGTTGGGG 912  
| | | | |  
QY 900 GAAAGACATTTAGACCTTTTAAAGTTACTGATTTAATTTTCATGTTATTTGGTTGCA 959  
| | | | |  
Db 913 GAAAGACATTTAGACCTTTTAAAGTTACTGATTTAATTTTCATGTTATTTGGTTGCA 972  
| | | | |  
QY 960 TGAAGTTGCCCTTAACCACTAAGGATTTATCAAGATTTTTCGCGAGACTTATACATGTCTA 1019  
| | | | |  
Db 973 TGAAGTTGCCCTTAACCACTAAGGATTTATCAAGATTTTTCGCGAGACTTATACATGTCTA 1032  
| | | | |  
QY 1020 GGATC 1024  
| | | | |  
Db 1033 GGATC 1037  
| | | | |

## RESULT 14

US-09-726-806-5292  
; Sequence 5292, Application US/09726806  
; GENERAL INFORMATION:  
; APPLICANT: Galvin, Katherine  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: 1600.2028-001  
; CURRENT APPLICATION NUMBER: US/09/726,806  
; CURRENT FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 60/168,135  
; PRIOR FILING DATE: 1999-11-30  
; NUMBER OF SEQ ID NOS: 6283  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5292  
; LENGTH: 2342  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-726-806-5292

Query Match 98.9%; Score 1013; DB 32; Length 2342;  
Best Local Similarity 99.9%; Pred. No. 1.6e-220;  
Matches 1024; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 1 ACCGCTCGTCCGCGCTTGAAGCGCGCGGAGCGC-GCGCAATTCGTCGCGCGCGG 59  
| | | | |  
Db 13 ACCGCTCGTCCGCGCTTGAAGCGCGCGGAGCGCGCAATTCGTCGCGCGCGG 72  
| | | | |

RESULT 15  
US-09-770-173-2543  
; Sequence 2543, Application US/09770173

QY 60 GGGGGCGGCTCCCGGCACTTTCCGCGCGCAACCAAGGACTACCAAGAGGGAGCGGCTGG 119  
| | | | |  
Db 73 GGGGGCGGCTCCCGGCACTTTCCGCGCGCAACCAAGGACTACCAAGAGGGAGCGGCTGG 132  
| | | | |  
QY 120 GATGGCGGTCCGCGCGCGCGGAGTACAAGCGGCGGAGCTTCTTCCGCAAGATGAA 179  
| | | | |  
Db 133 GATGGCGGTCCGCGCGCGCGGAGTACAAGCGGCGGAGCTTCTTCCGCAAGATGAA 192  
| | | | |  
QY 180 GGGCTACCCGCACTGCGCGCGCGGATTCAGTAACCTCCAGAGGGCGCTGTGAAGCTCC 239  
| | | | |  
Db 193 GGGCTACCCGCACTGCGCGCGCGGATTCAGTAACCTCCAGAGGGCGCTGTGAAGCTCC 252  
| | | | |  
QY 240 AGCAAAACAAGTATCTTCTTTTGGCACCCTATGAAACTGCACTTTCTAGGCTCCCA 299  
| | | | |  
Db 253 AGCAAAACAAGTATCTTCTTTTGGCACCCTATGAAACTGCACTTTCTAGGCTCCCA 312  
| | | | |  
QY 300 AGACCTTTTCCATATAAGAGAGTACAAGACAAGTTTGGAAAGTCAAAACGGAAGG 359  
| | | | |  
Db 313 AGACCTTTTCCATATAAGAGAGTACAAGACAAGTTTGGAAAGTCAAAACGGAAGG 372  
| | | | |  
QY 360 ATTTACAGAGGATTTGCGGGAATACAAATACCCAGGAGTAAAGTTTACTGGCTACCA 419  
| | | | |  
Db 373 ATTTACAGAGGATTTGCGGGAATACAAATACCCAGGAGTAAAGTTTACTGGCTACCA 432  
| | | | |  
QY 420 GGCNAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAAAGGTGGAAATACTGCAGATGC 479  
| | | | |  
Db 433 GGCNAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAAAGGTGGAAATACTGCAGATGC 492  
| | | | |  
QY 480 AAGCAGTGAAGAGAGGTGATAGTAGAAGAGATGGAAGAGGCAAAAGAAAGATGA 539  
| | | | |  
Db 493 AAGCAGTGAAGAGAGGTGATAGTAGAAGAGATGGAAGAGGCAAAAGAAAGATGA 552  
| | | | |  
QY 540 AAAAGCAGGCTCAAAACCGAAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCCG 599  
| | | | |  
Db 553 AAAAGCAGGCTCAAAACCGAAAAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCCG 612  
| | | | |  
QY 600 GAAATCTCCAGAGATGAAGATGACAAAGACTGCAAAAGAGAGAGAAACAAAAGCAGCTC 659  
| | | | |  
Db 613 GAAATCTCCAGAGATGAAGATGACAAAGACTGCAAAAGAGAGAGAAACAAAAGCAGCTC 672  
| | | | |  
QY 660 TGAGGTGGAGATGCGGGCAACGACACAAAGAAACAACTTCAGACTTCGAGAAAACCCAG 719  
| | | | |  
Db 673 TGAGGTGGAGATGCGGGCAACGACACAAAGAAACAACTTCAGACTTCGAGAAAACCCAG 732  
| | | | |  
QY 720 TGAAGGACCTAACTACCAATTAATGCTGCTGATATTAAGAGAAACCAAGAGAGTTAT 779  
| | | | |  
Db 733 TGAAGGACCTAACTACCAATTAATGCTGCTGATATTAAGAGAAACCAAGAGAGTTAT 792  
| | | | |  
QY 780 ATGTTGGTGTCTAATATCTTGGATTTGATATGAACCAACACATAGTCTCTTGTGTCA 839  
| | | | |  
Db 793 ATGTTGGTGTCTAATATCTTGGATTTGATATGAACCAACACATAGTCTCTTGTGTCA 852  
| | | | |  
QY 840 TTGACAGAACCCAGATTTGATGATATTAATTCCTCTCTGTTGTTGGGG 899  
| | | | |  
Db 853 TTGACAGAACCCAGATTTGATGATATTAATTCCTCTCTGTTGTTGGGG 912  
| | | | |  
QY 900 GAAAGACATTTAGACCTTTTAAAGTTACTGATTTAATTTTCATGTTATTTGGTTGCA 959  
| | | | |  
Db 913 GAAAGACATTTAGACCTTTTAAAGTTACTGATTTAATTTTCATGTTATTTGGTTGCA 972  
| | | | |  
QY 960 TGAAGTTGCCCTTAACCACTAAGGATTTATCAAGATTTTTCGCGAGACTTATACATGTCTA 1019  
| | | | |  
Db 973 TGAAGTTGCCCTTAACCACTAAGGATTTATCAAGATTTTTCGCGAGACTTATACATGTCTA 1032  
| | | | |  
QY 1020 GGATC 1024  
| | | | |  
Db 1033 GGATC 1037  
| | | | |

; GENERAL INFORMATION:  
; APPLICANT: White, David  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USBS  
; TITLE OF INVENTION: THEREFOR  
; FILE REFERENCE: 1600.2059-001  
; CURRENT APPLICATION NUMBER: US/09/770,173  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: US 60/178,876  
; PRIOR FILING DATE: 2000-01-28  
; NUMBER OF SEQ ID NOS: 3167  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2543  
; LENGTH: 1909  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-770-173-2543

Query Match 98.8%; Score 1012; DB 33; Length 1909;  
Best Local Similarity 99.9%; Pred. No. 2.5e-220;  
Matches 1023; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY	2	CGCTCGTCCGCGCGCTTGAGCGCGCGGAGGCGC-GCGCAATTCGTCGCGCGCGCGG 60
Db	34	CGCTCGTCCGCGCGCTTGAGCGCGCGCGGAGGCGCGCAATTCGTCGCGCGCGCGG 93
QY	61	GGGGCGGCTCCCGGCATCTTCGCGCGACCAAGGACTACCAAGGAGGGGAGCGGCTGG 120
Db	94	GGGGCGGCTCCCGGCATCTTCGCGCGACCAAGGACTACCAAGGAGGGGAGCGGCTGG 153
QY	121	ATGGGCGCTCCGCGCGCGCGAGTACAAAGCGGCGGCGCTTCGCGCAAGATGAAG 180
Db	154	ATGGGCGCTCCGCGCGCGCGAGTACAAAGCGGCGGCGCTTCGCGCAAGATGAAG 213
QY	181	GGCTACCGCACTGGCGCGCGCGGATGTAGTAACCTCCAGAGGCGCTGTGAAGCTCCA 240
Db	214	GGCTACCGCACTGGCGCGCGCGGATGTAGTAACCTCCAGAGGCGCTGTGAAGCTCCA 273
QY	241	GCAACACAGTATCCTATCTTCTTTTGGCACCCATGAACTGCACTTTCTAGTCCCAA 300
Db	274	GCAACACAGTATCCTATCTTCTTTTGGCACCCATGAACTGCACTTTCTAGTCCCAA 333
QY	301	GACCTTTTCCATATAAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAGGA 360
Db	334	GACCTTTTCCATATAAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAGGA 393
QY	361	TTTAACGAAGGATTTGTGGAAATAGAAATAACCCAGGAGTAAAGTTTACTGCGTACCAG 420
Db	394	TTTAACGAAGGATTTGTGGAAATAGAAATAACCCAGGAGTAAAGTTTACTGCGTACCAG 453
QY	421	GCAATTCAGCAACAGAGCTTTCAGAAACTGAGAGGAGAGGTGGAAATCTCAGATGCA 480
Db	454	GCAATTCAGCAACAGAGCTTTCAGAAACTGAGAGGAGAGGTGGAAATCTCAGATGCA 513
QY	481	AGCAGTGAGGAAGAGGTGATAGAGTAGAAGATGGAAGGCAAGGCAAGAAAGATGAA 540
Db	514	AGCAGTGAGGAAGAGGTGATAGAGTAGAAGATGGAAGGCAAGGCAAGAAAGATGAA 573
QY	541	AAAGCAGGCTCAAAACGGAAGGATGATATCTTCAAGAAATCTCTTAAACAGTCCCGG 600
Db	574	AAAGCAGGCTCAAAACGGAAGGATGATATCTTCAAGAAATCTCTTAAACAGTCCCGG 633
QY	601	AAATCTCCAGGAGTGAAGTACAAAGCTGCAAGAGAGGAAACAAAGAGCAGCTCT 660
Db	634	AAATCTCCAGGAGTGAAGTACAAAGCTGCAAGAGAGGAAACAAAGAGCAGCTCT 693
QY	661	GAGGCTGAGAGTGGGGCAACGACACAGAAACAACTTCAGACTTGCAGAAACCACT 720
Db	694	GAGGCTGAGAGTGGGGCAACGACACAGAAACAACTTCAGACTTGCAGAAACCACT 753
QY	721	GAAGGACCTTAACCTACCATATGAATGCTGCATATTTAGAGAAACCAACAGAGGTTATA 780
Db	754	GAAGGACCTTAACCTACCATATGAATGCTGCATATTTAGAGAAACCAACAGAGGTTATA 813

QY	781	TGTTTGGTTGCTAATATTTCTTGGATTTGATATGAACCAACACATAGTCTTGTGTCTAT 840
Db	814	TGTTTGGTTGCTAATATTTCTTGGATTTGATATGAACCAACACATAGTCTTGTGTCTAT 873
QY	841	TGACAGAAACCCAGTTTGTATGTACATTTATTCATATTCCTCTCTGTTGTGTTTCGGGGG 900
Db	874	TGACAGAAACCCAGTTTGTATGTACATTTATTCATATTCCTCTCTGTTGTGTTTCGGGGG 933
QY	901	AAAAGACATTTTACGCTTTTAAAGTTACTGATTTAAATTTCAATGTTTATTTGGTTGAT 960
Db	934	AAAAGACATTTTACGCTTTTAAAGTTACTGATTTAAATTTCAATGTTTATTTGGTTGAT 993
QY	961	GAAAGTTGCCCTTAACCACTTAAGGATTAACAAGATTTTGGCGAGACTTATACATGCTAG 1020
Db	994	GAAAGTTGCCCTTAACCACTTAAGGATTAACAAGATTTTGGCGAGACTTATACATGCTAG 1053
QY	1021	GATC 1024
Db	1054	GATC 1057

Search completed: January 18, 2004, 12:01:11  
Job time : 3670 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2004, 08:34:45 ; Search time 2527 Seconds  
(without alignments)  
9848.747 Million cell updates/sec

Title: US-09-787-328B-3  
Perfect score: 1024  
Sequence: 1 accgcgtccgcgcggctt.....acttatacatgttagatc 1024

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 22781392 seqs, 1215238056 residues  
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_estc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: gb\_est6:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pin:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1007.2	98.4	2327	11 BC040554	BC040554 Homo sapi
2	977.8	95.5	1201	9 AL520288	AL520288 AL520288
3	907	88.6	1201	9 AL520469	AL520469 AL520469
4	764	74.6	808	10 BG201738	BG201738 RST21080

5	763.2	74.5	1112	12	BM457409	BM457409
6	758	74.0	1868	11	AK044858	Mus muscu
7	757.8	74.0	889	14	CD251499	AGENCOURT
8	754.4	73.7	2722	11	AK019487	Mus muscu
9	746	72.9	1526	11	AK014849	Mus muscu
10	746	72.9	3363	11	AK045542	Mus muscu
11	739	72.2	819	10	BG204817	RST24232
12	690.4	67.4	692	12	BM722135	UI-B-EJO-
13	672.2	65.6	781	10	BE886964	601506447
14	659.6	64.4	813	10	BG212737	RST32332
15	653	63.8	796	14	CA315695	UI-M-FW0-
16	639.6	62.5	772	14	CD000122	AGENCOURT
17	631	61.6	643	14	CB215502	NISC np06
18	630.8	61.6	919	10	BF346836	602021638
19	623.6	60.9	737	9	AU126115	AU126115
20	613.2	59.9	914	10	BF306828	601891619
21	605.2	59.1	681	10	BG719650	602689854
22	597.6	58.4	664	12	B1400803	MI-P-AY1-
23	593.6	58.0	2788	11	AK051401	Mus muscu
24	576.8	56.3	671	14	BY721555	BY721555
25	572.8	55.9	707	12	B1202384	UI-M-CG09
26	568	55.5	863	13	BUI79069	AGENCOURT
27	562.6	54.9	876	13	BUI64599	AGENCOURT
28	562.4	54.9	580	9	AI743609	AI743609
29	552.8	54.0	557	14	CA941241	1r32602.x
30	551.8	53.9	651	14	CD354909	UI-M-GM0-
31	549.4	53.7	663	10	BE111763	UI-R-BJ1-
32	543	53.0	559	10	BE703369	MR2-NN111
33	539	52.6	566	10	BE703403	MR2-NN111
34	536.2	52.4	654	10	BE703421	MR2-NN111
35	527.2	51.5	532	10	BF445986	7p17d09.x
36	525.6	51.3	557	10	BE348635	htc73c04.x
37	525.4	51.3	547	12	BM931689	UI-B-EJ1-
38	522.8	51.1	973	13	BQ685483	AGENCOURT
39	522.4	51.0	545	12	BM931237	UI-B-EJ1-
40	522.4	51.0	545	13	BQ186294	UI-B-EJ1-
41	522.4	51.0	560	10	BE550171	7b49g08.x
42	519.6	50.7	538	12	BM725299	UI-B-EJ0-
43	514.2	50.2	531	9	AW161910	au71e02.x
44	513.6	50.2	608	10	BE703418	MR2-NN111
45	513	50.1	602	10	BE703252	MR2-NN111

ALIGNMENTS

RESULT 1  
BC040554

LOCUS  
DEFINITION

Homo sapiens, similar to likely ortholog of mouse hepatoma-derived growth factor, related protein 3, clone IMAGE:5303793, mRNA.

ACCESSION

BC040554

VERSION

BC040554.1

KEYWORDS

HTC

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 2327)

AUTHORS

Strausberg, R.

TITLE

Direct Submission

JOURNAL

Submitted (29-NOV-2002)

REMARK

NIH-MGC Project URL: http://mgc.nci.nih.gov

COMMENT

Contact: MGC help desk

Email: cgaops@mail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki

Toshiyuki and Piero Carninci (RIKEN)

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome

QY	2	CCGCTCGTCGCGCCCGGCTTGAGGCCCGCGGGAGCGC-GGCAATTCGTGGCCCGCGGG	60
DB	488	CCGCTCGTCGCGCCCGGCTTGAGGCCCGCGGGAGCGGGCGCAATTCGTGGCCCGCGGG	547
QY	61	GGGCGGCCCTCCGCGCATCTTCGGCGGCGACCAAGAGACTACAGGAAGGGAGCGGCTGGG	120
DB	548	GGGCGGCCCTCCGCGCATCTTCACGGCGACCAAGGACTACAGGAAGGGAGCGGCTGGG	607
QY	1.21	ATGCGCGCTCCGCGGCCCGCGAGTCAAAAGCGGCGACCTGGTCTTCGCGCAAGATGAAG	180
DB	608	ATGCGCGCTCCGCGGCCCAACGAGTCAAAAGCGGCGACCTGGTCTTCGCGCAAGATGAAG	667
QY	181	GGGTACCGGCACCTGGCGCGGCCGGATGTAGTAATCCAGAGGGCGCTGTGAAGCTCCA	240
DB	668	GGGTACCGGCACCTGGCGCGGCCGGATGTAGTAATCCAGAGGGCGCTGTGAAGCTCCA	727
QY	241	GCAAAACAAGTATCCTATCTCTCTTTTGGCACCCATGAAACTGGATTTCTAGGTGCCAAA	300
DB	728	GCAAAACAAGTATCCTATCTCTCTTTTGGCACCCATGAAACTGGATTTCTAGGTGCCAAA	787
QY	301	GACCTTTTCCATATAAGGAGTCAAAAGACAAGTTTGGAAAGTCAAAACAAACGGAAGGA	360
DB	788	GACCTTTTCCATATAAGGAGTCAAAAGACAAGTTTGGAAAGTCAAAACAAACGGAAGGA	847
QY	361	TTTAAACGAAGGATTGTGGGAAATAGAAAATAACCCAGAGGTAAAGTTTACTGGCTACAG	420
DB	848	TTTAAACGAAGGATTGTGGGAAATAGAAAATAACCCAGAGGTAAAGTTTACTGGCTACAG	907
QY	421	GCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAGGTGGAATACTGCAGATGCA	480
DB	908	GCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAGGTGGAATACTGCAGATGCA	967
QY	481	AGCAGTCAGGAAGAAGGTGATAGAGTAGAAGAAGATGGAAGGCAAAAGAAAGAAATGAA	540
DB	968	AGCAGTCAGGAAGAAGGTGATAGAGTAGAAGAAGATGGAAGGCAAAAGAAAGAAATGAA	1027
QY	541	AAAGCAGGCTCAAAACGGAAAAGTCAATATCTTCAAGAAAATCCCTCTAAACAGTCCCGG	600
DB	1028	AAAGCAGGCTCAAAACGGAAAAGTCAATATCTTCAAGAAAATCCCTCTAAACAGTCCCGG	1087
QY	601	AAATCTCCAGGAGTGAAGATGACAAAGACTGCAAGAGAGGAAAACAAAAGCAGCTCT	660
DB	1088	AAATCTCCAGGAGTGAAGATGACAAAGACTGCAAGAGAGGAAAACAAAAGCAGCTCT	1147
QY	661	GAGGGTGGAGATCGGGCAACGACACAAGAAAACAATCTTCAGACTTGCAGAAAACCAAGT	720

Matches 1013; Conservative 1; Mismatches 3; Indels 4; Gaps 3;			
Qy	5	CTGTCGCCCGCGGCTTGAGGCGCGCGGGAGCGC-GGCAATTCGTGCGCCCGCGGGGG	63
Db	66	CTGTCGCCCGCGGCTTGAGGCGCGCGGGAGCGGCAATTCGTGCGCCCGCGGGGG	125
Qy	64	GCGGCTCCCGGCATCTTCGCGGCGACCAAGGACTACAGAGGGGAGCGGTGGGATG	123
Db	126	GCGGCTCCCGGCATCTTCGCGGCGACCAAGGACTACAGAGGGGAGCGGTGGGATG	185
Qy	124	GCGGCTCCCGGCGCGCGAGTACAAAGCGGGCGACCTGCTTCGCCAAGATCAAGGC	183
Db	186	GCGGCTCCCGGCGCGCGAGTACAAAGCGGGCGACCTGCTTCGCCAAGATCAAGGC	245
Qy	184	TACCGCACTGGCGGCGCGGATGATGAATCCAGAGGGCGCTGTGAAGCTCCAGCA	243
Db	246	TACCGCACTGGCGGCGCGGATGATGAATCCAGAGGGCGCTGTGAAGCTCCAGCA	305
Qy	244	AACAGTATCTATCTCTCTTTTGGCACCCATGAATCTGATTTCTAGTCTCCAAAGAC	303
Db	306	AACAGTATCTATCTCTCTTTTGGCACCCATGAATCTGATTTCTAGTCTCCAAAGAC	365
Qy	304	CTTTTCCATATAGGAGTCAAGACAAGTTTGGAAAGTCMAACACGAAGGATTT	363
Db	366	CTTTTCCATATAGGAGTCAAGACAAGTTTGGAAAGTCMAACACGAAGGATTT	425
Qy	364	AACGAAGATGTGGGAAATAGAAATAAACCCAGGAGTAAAGTTTACTGGCTACCAAGCA	423
Db	426	AACGAAGATGTGGGAAATAGAAATAAACCCAGGAGTAAAGTTTACTGGCTACCAAGCA	485
Qy	424	ATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAGGTGGAATATCTGCAGATGCAAGC	483
Db	486	ATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAGGTGGAATATCTGCAGATGCAAGC	545
Qy	484	AGTGAGGAAGAGTGATAGTAGTAGAGAGATGGAAGGCAAAAGCAAGATGAATA	543
Db	546	AGTGAGGAAGAGTGATAGTAGTAGAGAGATGGAAGGCAAAAGCAAGATGAATA	605
Qy	544	GCAGGCTCAAAACGGGAAAGTATATATCTTCAAGAAATCTCTTAAACAGTCCCGGAA	603
Db	606	GCAGGCTCAAAACGGGAAAGTATATATCTTCAAGAAATCTCTTAAACAGTCCCGGAA	665
Qy	604	TCTCCAGGAGATGAAGATGACAAAGCTGCAAGAGAGAGGAAACAAACAGCAGCTGAG	663
Db	666	TCTCCAGGAGATGAAGATGACAAAGCTGCAAGAGAGAGGAAACAAACAGCAGCTGAG	725
Qy	664	GTTGGAGATGCGGCAACGACACAGAAACACAACTTCAGACTTGCAGAAACAGTGAA	723
Db	726	GTTGGAGATGCGGCAACGACACAGAAACACAACTTCAGACTTGCAGAAACAGTGAA	785
Qy	724	GGGACCTTAATACCATATGAATGCTGCATATTAAGAGAAACCAAGAAAGGTTATATGT	783
Db	786	GGGACCTTAATACCATATGAATGCTGCATATTAAGAGAAACCAAGAAAGGTTATATGT	845
Qy	784	TGTTGTTCTTAATATCTTGGATTTGATATGAACCAACACATAGTCTCTTGTTCATTTGA	843
Db	846	TGTTGTTCTTAATATCTTGGATTTGATATGAACCAACACATAGTCTCTTGTTCATTTGA	905
Qy	844	CAGAACCCAGTTTGATGTACATATTCATATTCCTCTCTGTGTGTTTCGGGGGAA	903
Db	906	CAGAACCCAGTTTGATGTACATATTCATATTCCTCTCTGTGTGTTTCGGGGGAA	965
Qy	904	AGACATTTTACCTTTTAAAGTACTGATTAATTTTCAATGTTATTTTGGTTCATGAA	963
Db	966	AGACATTTTACCTTTTAAAGTACTGATTAATTTTCAATGTTATTTTGGTTCATGAA	1025
Qy	964	GTTGCCCTTAACCACTAAGGATATCAAGATTTTGGCGAGACTTATACATGTCTAGGAT	1023
Db	1026	GTTGCC--TTAACACTAAGGATATCAAGATTTTGGCG--GACTTATACATGTCTAGGAT	1082
Qy	1024	C 1024	
Db	1083	C 1083	

RESULT 3

AL520469

LOCUS

DEFINITION

AL520469

ACCESSION

AL520469

VERSION

AL520469.2

KEYWORDS

EST.

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1201)

Li, W. B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

On Feb 13, 2001 this sequence version replaced gi:12783962.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 7531.r For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DB006QPI&cluster=7531.r. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/

Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DB006BD06QPI.

Location/Qualifiers

1. .1201

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DB006YG12"

/issue\_type="NEUROBLASTOMA COT 10-NORMALIZED"

/clone\_lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 378 a 240 c 299 g 260 t 24 others

ORIGIN

Query Match 88.6%; Score 907; DB 9; Length 1201;

Best Local Similarity 98.9%; Pred. No. 9.9e-153;

Matches 954; Conservative 1; Mismatches 6; Indels 4; Gaps 4;

Qy 10 CCGCCCGGCTTGAGGCGCGCGGGAGCGC-GGCAATTCGTGCGCCCGCGGGGGCGGC 68

Db 62 CCGGATGCTTGAGGCGCGCGGGAGCGGCAATTCGTGCGCCCGCGGGGGGGCGGC 121

Qy 69 CTCGCCGATCTTTCGCGGCGACCAAGGACTACAGAGGGAGCGGCTGGGATGGCGG 128

Db 122 CTCGCCGATCTTTCGCGGCGACCAAGGACTACAGAGGGAGCGGCTGGGATGGCGG 181

Qy 129 TCGCGGGCGCGCGAGTACAAAGCGGGCGACCTGTGCTTCGCCAAGATGAGGGCTACCC 188

Db 182 TCGCGGGCGCGCGAGTACAAAGCGGGCGACCTGTGCTTCGCCAAGATGAGGGCTACCC 241

Qy 189 GCACTGGCGGGCGCGGATTTGATGAACTCCAGAGGGCGCTGTGAAGCTCCAGCAACAA 248

Db 242 GCACTGGCGGGCGCGGATTTGATGAACTCCAGAGGGCGCTGTGAAGCTCCAGCAACAA 301

Qy 249 GTATCTTATCTTCTTTTGGCACCCCATGAACTGCAATTTCTAGGTCCCAAGACCTTTT 308

Db 302 GTATCTTATCTTCTTTTGGCACCCCATGAACTGCAATTTCTAGGTCCCAAGACCTTTT 361

Qy 309 TCATATAAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAGGATTTAACGA 368

Db 362 TCATATAAGGAGTACAAAGCAAGTTTGGAAAGTCAAAACAAACGGAAGGATTTAACGA 421

QY 369 AGGATTGTGGAAATAGAAATAAACCAGAGTAAAGTTTACTGTGCTACCGCAATTCA 428  
 Db 422 AGGATTGTGGAAATAGAAATAAACCAGAGTAAAGTTTACTGTGCTACCGCAATTCA 481  
 QY 429 GCAACAGAGCTCTTCAGAACTGAGGAGAGGTGAAATCTGCAGATGCAAGCAGTGA 488  
 Db 482 GCAACAGAGCTCTTCAGAACTGAGGAGAGGTGAAATCTGCAGATGCAAGCAGTGA 541  
 QY 489 GGAAGAGGTGATAGAGTGAAGAGATGGAAGGCAAAAGAAAGATGAAAGACGAGG 548  
 Db 542 GGAAGAGGTGATAGAGTGAAGAGATGGAAGGCAAAAGAAAGATGAAAGACGAGG 601  
 QY 549 CTCAAAACGGAAAAAGTCTATATCTTCAAGAAATCTCTTAACAGTCCCGAAATCTCC 608  
 Db 602 CTCAAAACGGAAAAAGTCTATATCTTCAAGAAATCTCTTAACAGTCCCGAAATCTCC 661  
 QY 609 AGGAGATGAAGATGACAAAGACTGCAAGAGAGGAGAAACAAAGCAGCTCTGAGGGTGG 668  
 Db 662 AGGAGATGAAGATGACAAAGACTGCAAGAGAGGAGAAACAAAGCAGCTCTGAGGGTGG 721  
 QY 669 AGATGGGGCAACGACACAAGAAACACAACTTCAGACTTTCAGAGAAACCAAGTGAAGGAC 728  
 Db 722 AGATGGGGCAACGACACAAGAAACACAACTTCAGACTTTCAGAGAAACCAAGTGAAGGAC 781  
 QY 729 CTAACCTACCATATGAATGCTGCATATTAAGAGAAACCAAGAGAGTTATATGTTGT 788  
 Db 782 CTAACCTACCATATGAATGCTGCATATTAAGAGAAACCAAGAGAGTTATATGTTGT 841  
 QY 789 TGCTCTAATATCTTGGAATTTGATATGAACCAACACATAGTCTCTGTTGTCATTCAGAAA 848  
 Db 842 TGCTCTAATATCTTGGAATTTGATATGAACCAACACATAGTCTCTGTTGTCATTCAGAAA 901  
 QY 849 CCCAGTTTGTATGATATATTCCTCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 908  
 Db 902 -CCCAGTTTGTATGATATATTCCTCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 959  
 QY 909 TTTTAGCCCTTTTAAAGTACTGATTTAAATTTTCTGTTGTTGTTGTTGTTGTTGTTGTTG 968  
 Db 960 TTTTAGCCCTTTTAAAGTACTGATTTAAATTTTCTGTTGTTGTTGTTGTTGTTGTTGTTG 1018  
 QY 969 CCTTA 973  
 Db 1019 CTATA 1023

RESULT 4  
 BG201738  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 Harrington, J.J., Sherif, B., Rundlett, S., Jackson, P.D., Perry, R.,  
 Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J.,  
 Lerner, L., Costanzo, D., McElligott, K., Booser, S., Mays, R., Smith,  
 E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K., Offenbacher,  
 J., Danzig, J., and Ducar, M.  
 Creation of genome-wide protein expression libraries using random  
 activation of gene expression  
 Nat. Biotechnol. 19 (5), 440-445 (2001)  
 21227151  
 PUBMED  
 11329013  
 COMMENT  
 Contact: Scott J. Cain  
 Athersys, Inc.  
 3201 Carnegie Ave., Cleveland, OH 44115, USA  
 Tel: 216 431 9900  
 Fax: 216 361 9596

Email: scain@atersys.com  
 High quality sequence stop: 511.  
 Location/Qualifiers  
 1..808  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /cell\_line="HT1080"  
 /clone\_lib="Athersys RAGE Library"  
 /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."  
 292 a 141 c 181 g 194 t

FEATURES  
 source

Query Match 74.6%; Score 764; DB 10; Length 808;  
 Best Local Similarity 99.0%; Pred. No. 4.5e-127;  
 Matches 800; Conservative 0; Mismatches 5; Indels 3; Gaps 3;  
 QY 205 ATTGATGAATCCCGAGGGCGCTGTGAAGCTCCAGCAAAACAATATCTCTTTT 264  
 Db 1 ATTGATGAATCCCGAGGGCGCTGTGAAGCTCCAGCAAAACAATATCTCTTTT 60  
 QY 265 TTTGGCACCCATGAACCTGCAATTTCTAGGTCCTCAAGACCTTTTCCATATAAGGAGTAC 324  
 Db 61 TTTGGCACCCATGAACCTGCAATTTCTAGGTCCTCAAGACCTTTTCCATATAAGGAGTAC 120  
 QY 325 AAAGACAAGTTTGGAAAGTCAAAACAAACGAAAGATTTAAACGAAGGATTTGTGGAAATA 384  
 Db 121 AAAGACAAGTTTGGAAAGTCAAAACAAACGAAAGATTTAAACGAAGGATTTGTGGAAATA 180  
 QY 385 GAAATAATACCAGAGTAAAGTTTCTGGCTACAGGCAATTCAGCAACAGAGCTCTTCA 444  
 Db 181 GAAATAATACCAGAGTAAAGTTTCTGGCTACAGGCAATTCAGCAACAGAGCTCTTCA 240  
 QY 445 GAACTGAGGAGAGGTGGAATTAATCTGACATCCAGCAGTGGAGAGAGGTGATAGA 504  
 Db 241 GAACTGAGGAGAGGTGGAATTAATCTGACATCCAGCAGTGGAGAGAGGTGATAGA 300  
 QY 505 GTAGAAGAGATGGAAGGCAAAAGAAAGATGAAAGCAGGCTCAAAACGGAAAAAG 564  
 Db 301 GTAGAAGAGATGGAAGGCAAAAGAAAGATGAAAGCAGGCTCAAAACGGAAAAAG 360  
 QY 565 TCATATCTTCAAGAAATCTCTTAAACAGTCCCGGAAATCTCCAGGAGATCAAGATGAC 624  
 Db 361 TCATATCTTCAAGAAATCTCTTAAACAGTCCCGGAAATCTCCAGGAGATCAAGATGAC 420  
 QY 625 AAAGACTGCAAGAGAGGAAACAAAGAGCAGCTCTGAGGGTGGAGATCGGGCAACGAC 684  
 Db 421 AAAGACTGCAAGAGAGGAAACAAAGAGCAGCTCTGAGGGTGGAGATCGGGCAACGAC 480  
 QY 685 ACAAGAAACACAACCTTCAGACTTCGAGAAAAACCAAGTGAAGGACCTTAACATAATGA 744  
 Db 481 ACAAGAAACACAACCTTCAGACTTCGAGAAAAACCAAGTGAAGGACCTTAACATAATGA 540  
 QY 745 ATGCTGCATATTAAGAGAAACCAAGAAAGTTATATCTTTGGTTGTTCTAATATCTTGG 804  
 Db 541 ATGCTGCATATTAAGAGAAACCAAGAAAGTTATATCTTTGGTTGTTCTAATATCTTGG 600  
 QY 805 ATTTGATATGACCAACACATAGTCTTGTGTCATTCAGACAGACCCAGTTGTATGTA 864  
 Db 601 ATTTGATATGACCAACACATAGTCTTGTGTCATTCAGACAGACCCAGTTGTATGTA 660  
 QY 865 CATTATTCATATCTCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 924  
 Db 661 CATTATTCATATCTCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 718  
 QY 925 AAGTTACTGATTAATTTTCATGTTTATTTGGTTGATGAAGTTGCCCTTAAACCACTAAGA 984  
 Db 719 AAGTTACTGATTAATTTTCATGTTTATTTGGGTGATGAAGTTGCCCTTAAACCACTAAGA 777

```

Qy 985 TTATCAAGATTTTTCGCGAGACTTATAC 1012
Db 778 TTATCAAGATTTTTCGCGAGACTTATAC 805

RESULT 5
BM457409 1112 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT 6407463 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5584879
DEFINITION 5', mRNA sequence.
ACCESSION BM457409
VERSION BM457409.1 GI:18506449
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1112)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12350 Row: e Column: 08
High quality sequence stop: 541.
FEATURES
source
1..1112
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5584879"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 385 a 211 c 228 g 288 t
ORIGIN
Query Match 74.5%; Score 763.2; DB 12; Length 1112;
Best Local Similarity 98.9%; Pred. No. 5.9e-127;
Matches 779; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Qy 236 CTCGAGCAACAAGATTCCTATCTCTCTTTTGGCACCCATGAACATGCAATTTCTAGGTC 295
Db 1 CTCGAGCAACAAGATTCCTATCTCTCTTTTGGCACCCATGAACATGCAATTTCTAGGTC 60

Qy 296 CCAAGACCTTTTCCATATAAGAGTACAAAGCAAGTTTGGAAAGTCAACAAACGGA 355
Db 61 CCAAGACCTTTTCCATATAAGAGTACAAAGCAAGTTTGGAAAGTCAACAAACGGA 120

Qy 356 AAGGATTTAAGGAGGTTGGGAAATAGAAATACCCAGGAGTAAAGTTTACTGGCT 415
Db 121 AAGGATTTAAGGAGGTTGGGAAATAGAAATACCCAGGAGTAAAGTTTACTGGCT 180

Qy 416 ACCAGGCAATTCAGCAACAGAGCTCTTCAGAAATCTGAGGAGAGGTGGAAATCTGCAG 475
Db 181 ACCAGGCAATTCAGCAACAGAGCTCTTCAGAAATCTGAGGAGAGGTGGAAATCTGCAG 240

Qy 476 ATGCAAGCAGTGGAGGAAGGTGATAGAGTAGAAGAGTGGAAAGGCAAGAAAGA 535
Db 241 ATGCAAGCAGTGGAGGAAGGTGATAGAGTAGAAGAGTGGAAAGGCAAGAAAGA 300

```

---

```

Qy 536 ATGAAAAAGCAGGCTCAAAACGGAAAAAGTCATATACTTCAAGAAAAATCCTCTAAACAGT 595
Db 301 ATGAAAAAGCAGGCTCAAAACGGAAAAAGTCATATACTTCAAGAAAAATCCTCTAAACAGT 360

Qy 596 CCCGAAATCTCCAGGAGATGAAGATGCAAAAGCTGCAAAAGAGGAGGAAAAACAAAGCA 655
Db 361 CCCGAAATCTCCAGGAGATGAAGATGCAAAAGCTGCAAAAGAGGAGGAAAAACAAAGCA 420

Qy 656 GCTCTGAGGCTGAGAGATCGGGCAACGACACACAGAAACCAACTTCAGACTTCGACAGAAA 715
Db 421 GCTCTGAGGCTGAGAGATCGGGCAACGACACACAGAAACCAACTTCAGACTTCGACAGAAA 480

Qy 716 CCAGTGAAGGGACCTAACTACCAATAATGAATGCTGCATATTTAAGAGAAAAACCAAGAGG 775
Db 481 CCAGTGAAGGGACCTAACTACCAATAATGAATGCTGCATATTTAAGAGAAAAACCAAGAGG 540

Qy 776 TTATATGTTTGGTGTCTAATATCTTGGATTTGATATGAAACCAACATAGTCCTTGT 835
Db 541 TTATATGTTTGGTGTCTAATATCTTGGATTTGATATGAAACCAACATAGTCCTTGT 600

Qy 836 GTCAATTGACAGAACCCAGCTTTGTATGTACATTTATCATATTCCTCTCTGTCGTGTT-C 894
Db 601 GTCAATTGACAGAACCCAGCTTTGTATGTACATTTATCATATTCCTCTCTGTCGTGTTCC 660

Qy 895 GGGGGGAAAAAGACATTTTACGCCCTTTTAAAGGTTACTGATTTAATTTTCATGTTATTTGG 954
Db 661 GGGGGGAAAAAGACATTTTACGCCCTTTTAAAGGTTACTGATTTAATTTTCATGTTATTTGG 720

Qy 955 TTGCATGAAGTTGCCCTTAAACCACTAAGGATTAATCAAGATTTTTCGCGAGACTTATACAT 1014
Db 721 GTGCTGAAAGTTGCCCTTAAACCACTAAGGATTAATCAAGATTTTTCGCGAGACTTATACAT 780

Qy 1015 GTCTAGGA 1022
Db 781 GTCTAGGA 788

RESULT 6
AK044858 1868 bp mRNA linear HTC 05-DEC-2002
LOCUS Mus musculus 9.5 days embryo parthenogenote cDNA, RIKEN full-length
DEFINITION enriched library, clone: B130008G01 product: hepatoma-derived growth
factor, related protein 3, full insert sequence.
ACCESSION AK044858
VERSION AK044858.1 GI:26336878
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE Carninci, P. and Hayashizaki, Y.
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

```

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861  
4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G.,  
Quackenbush, J., Schriml, L. M., Staubli, P., Suzuki, R., Tomita, M.,  
Wagner, I., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
King, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,  
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.  
and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 1868)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.

Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Tissues were provided by Dr. Tomohiro Kono (Department of Animal  
Science, Tokyo University of Agriculture, 1737 Hunko Atsugi City,  
Kanagawa Prefecture, Japan) whose assistance we gratefully  
acknowledge.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/.

Location/Qualifiers  
1. .1868  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"

FEATURES  
source

/db\_xref="FANTOM\_DB:B130008G01"  
/db\_xref="taxon:10090"  
/clone="B130008G01"  
/tissue\_type="parthenogenote"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="9.5 days embryo"  
263. .871  
/note="unnamed protein product; hepatoma-derived growth  
factor, related protein 3 (MGD|MG1:1352760, GB|NM\_013886,  
evidence: BLASTN, 99%, match=611)  
putative"  
/codon\_start=1  
/protein\_id="BAC32123.1"  
/db\_xref="GI:26336879"  
/translation="MARPRPREYKAGDLVFAKMGYPHPARIDELPGAVKPPANKY  
PIFFFGTHETAFGLPKDLPFYKEYDKFKGNKRNKGFNEGWEIENNVGRTGTQTI  
QQSSSTEGGNTADASSEEDRVEDGKRNKKGSKRSKYSYTKSSKQSRK  
SPGDEDDKCKEENKSSSEGGAGNDRNTTADLQAGEGT"  
BASE COUNT 519 a 414 c 457 g 478 t

## ORIGIN

Query Match 74.0%; Score 758; DB 11; Length 1868;  
Best Local Similarity 86.5%; Pred. No. 4.5e-126;  
Matches 887; Conservative 0; Mismatches 125; Indels 14; Gaps 4;  
QY 2 CGGCTCGTCCCGCGCTTGAGGCC---CGCGGGAGCGCGCGCAATTCGTGGCGCGCG 58  
DB 141 CGGCTCGAGCGTTCGGCTCGAGAGCGCGCGCGCGCGCGTTCGCTGTCGCGCG 200  
QY 59 GGGGGCGCGCTCCCGCGCATCTTCGCGCGCAACAGAGACTACCAGAGAGGGAGCGGCTG 118  
DB 201 GAGGGCGCGCTCCGAGCATCTTCGCGCGCGCAAGGACCGCGGAGGAGAGACGCG 260  
QY 119 GGATGGCGGTCCGCGCGCGCGAGTACAAAGCGGGCGACCTGCTTTCGCGCAAGATGA 178  
DB 261 GGATGGCGGTCCGCGCGCGCGAGTACAAAGCGGGAGACTGCTGCTTTCGCGCAAGATGA 320  
QY 179 AGGGTACCCGACTGCGCGCGCGGATGATGAATCTCCAGAGGGCGCTGTGAAGCCTC 238  
DB 321 AGGGTACCCGACTGCGCGCGCGGATGATGAATCTCCAGAGGGAGCTGTGAAGCCTC 380  
QY 239 CAGCAAAACAAGTATCTATCTCTTTTGGCACCATCAAACTGCAATTTCTAGTCCCA 298  
DB 381 CAGCAAAACAAGTATCTATCTCTTTTGGTACCCATGAACTGCAATTTCTAGTCCCA 440  
QY 299 AAGACCTTTTCCATATAAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACGGAAG 358  
DB 441 AAGACCTTTTCCATATAAGGAGTACAAAGACAAGTTTGGAAAGTCAAAACGGAAG 500  
QY 359 GATTTACGAGGATCTGGGGAATAGAAATACCCAGGAGTAAAGTTTACTGGCTACC 418  
DB 501 GATTTAATGAAGGATTTATGGGAAATTTGAAATATATCCAGGAGTAAATTTACTGGGTACC 560  
QY 419 AGGCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAGAGTGGAAATACTGCAGATG 478  
DB 561 AGCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGGAGAGAGTGGAAATACTGCAGATG 620  
QY 479 CAACAGTGAAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 538  
DB 621 CAACAGTGAAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 677  
QY 539 AAAAAGCAGGCTCAAAACGGGAAAGTCATATCTTCAAGAAATCTCTTAAACAGTCCC 598  
DB 678 AAAAAGGAGGCTCAAAACGGGAAAGTCTTACAGTCAAGAAAGTCTTCTTAAACAGTCCC 737  
QY 599 GGAATCTCCAGGAGATCAAGATCAAGACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGT 658  
DB 738 GGAATCTCCAGGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 797  
QY 659 CTGAGGGTGAAGATGCGGGCAACGACACAAAGAAACAACTTTCAGACTTTCAGAGAAACCA 718  
DB 798 CTGAGGGCGGAGACGCTGGCAATGACACGAGAAACACGACTTGCAGACTTTCGAGAAAGCGC 857









Db 185 ATTGATGAACCTCCAGAGGAGCTGTGAAGCTCCAGCAAAACAGTACCTATCTCTTT 244

Qy 265 TTGGCCACCCATGAATCTAGCTTCTAGGTCCCAAGACCTTTTCCATATAAGGAGTAC 324

Db 245 TTGGTACCCATGAATCTAGCTTCTAGGTCCCTAAGACCTTTTCCATATAAGGATAC 304

Qy 325 AAGACAAAGTTTGGAAAGTCAACAAACCGAAAGGATTTAAACGAAGATTGTGGGAAATA 384

Db 305 AAGACAAAGTTTGGAAAGTCAACAAACCGAAAGGATTTAAACGAAGATTGTGGGAAAT 364

Qy 385 GAAATAAACCCAGAGTAAAGTTTACTGCTTCCAGCAATTCAGCAACAGAGCTCTTCA 444

Db 365 GAAATAAATCCAGAGTAAAGTTTACTGCTTCCAGCAATTCAGCAACAGAGCTCTTCA 424

Qy 445 GAACTGAGGAGAGAGTGGAAATCTGCGAGTCAAGCAGTGAAGGAGAGGATGATAGA 504

Db 425 GAACTGAGGAGAGAGTGGAAATCTGCGAGTCAAGCAGTGAAGGAGAGGATGATAGA 484

Qy 505 GTAGAGAAAGTGAAGGAG 564

Db 485 GT---AGAGATGGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 541

Qy 565 TCATATCTTCAAGAAATCTCTTAACAGTCCCGGAATCTCCAGAGATGAAGATGAC 624

Db 542 TCCTACAGTCAAGAGAGTCTCTTAACAGTCCCGGAATCTCCAGAGATGAAGATGAT 601

Qy 625 AAAGACTGCAAG 684

Db 602 AAAGACTGCAAG 661

Qy 685 ACAAGAAACAACATCTCAGACTTGCAGAAACAGAGTGAAGGAGAGAGAGAGAGAGAG 744

Db 662 ACAGAGAAACAGACTGAGAGTGCAGAAACAGAGTGAAGGAGAGAGAGAGAGAGAGAG 721

Qy 745 ATCTGCTATTAAGAGAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 804

Db 722 ATCTGCTATTAAGAGAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 781

Qy 805 ATTTGATATGAACCAACATAGTCTCTGTTGTCATTGACAGAGAGAGAGAGAGAGAG 864

Db 782 TTGATTTGATATGAACCAACAGTCTCTGTTGTCATTGACAGAGAGAGAGAGAGAGAG 841

Qy 865 CATATTCATATCTCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 924

Db 842 CATATTCATATCTCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 900

Qy 925 AAGTACTGATTAATTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 984

Db 901 ----GGTTGATTTATTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 956

Qy 985 TTATCAAGATTTTGGCAGACTTATACATGCTCTAGATC 1024

Db 957 TGATGAAGATTTT---GCAGACTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 993

RESULT 9

AK014849

LOCUS

DEFINITION

1526 bp mRNA linear HTC 05-DEC-2002

Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4921509B08 product:hepatoma-derived growth factor, related protein 3, full insert sequence.

AK014849

ACCESSION

AK014849.1 GI:12852929

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

REFERENCE

1

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

AUTHORS

Carninci, P. and Hayashizaki, Y.

TITLE

High-efficiency full-length cDNA cloning

JOURNAL

Meth. Enzymol. 303, 19-44 (1999)

MEDLINE

99279253

10349636

REFERENCE  
AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

10349636

REFERENCE  
AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

10349636

REFERENCE  
AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

10349636

REFERENCE  
AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## PUBMED

## COMMENT

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5', GAGAGAGAGAGGATCAAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGCTGATTAATTAATTAATCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. cDNA of size comprised longer than 7 kb was selected before cloning. Vector: a modified pBluescript KS(+) after bulk excision from Lambda PLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

## FEATURES

Location/Qualifiers  
1..1526  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:4921509B08"  
/db\_xref="MGI:1894584"  
/db\_xref="taxon:10090"  
/clone="4921509B08"  
/sex="male"  
/tissue type="testis"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stages="adult"  
263..870

misc\_feature  
polya\_signal  
polya\_site

BASE COUNT 430 a 350 c 383 g 363 t

ORIGIN

Query Match 72.9%; Score 746; DB 11; Length 1526;  
Best Local Similarity 86.4%; Pred. No. 6.7e-124;  
Matches 886; Conservative 0; Mismatches 125; Indels 15; Gaps 5;

2 CCGCTGCTCCCGCCGCTTGAGGCC---CGCGGGAGCGCGCAATTGCTGCGCCCGCG 58  
141 CCGCTCGAGCGTTCGGCTCGAGGCCCGCGGGAGCGCGCGCTCGCTCGTCCCGCG 200  
59 GGGGGGGCGCTCCGGGATCTTCGGCGGACCAAGGACTACCAAGAGGGGAGCGGCTG 118  
201 GAGGGGGCGCTCCCGAGCATCTTCGGCGGCGCAAGGACCGCGGAGGAGGAGCGCG 260  
119 GGATGCGCGCTCCGGCGGCGCCCGAGTACAAAGCGGGCGACCTGGCTTCGCGCAAGATGA 178  
261 GGATGCGCGCTCCGGCGGCGCCCGAGTACAAAGCGGGAGCTGCTTCGCGCAAGATGA 320  
179 AGGGCTACCCGCACTCGCGCGCGCCGAGTGAATCCAGAGGGCGCTGTGAAGCCTC 238  
321 AGGGCTACCCGCACTCGCGCGCGCCGAGTGAATCCAGAGGGAGCTGTGAAGCCTC 380  
239 CAGCAACCAAGTATCCTATCTCTTTTGGCACCAGTGAATCTGATTTCTAGTCCCA 298  
381 CAGCAACCAAGTATCCTATCTCTTTTGGTACCAGTGAATCTGATTTCTAGTCCCA 440  
299 AAGACTTTTCCATATAGGAGTACAAAGCAAGTTTGGAAAGTCAACCAACCGGAAAG 358  
441 AAGACTTTTCCATATAGGATACAAAGCAAGTTTGGAAAGTCAACCAACCGGAAAG 500  
359 GATTTAAGCAAGATTTGGGAAATAGAAATAAACCCAGGAGTAAGTTTACTGCTACC 418  
501 GATTTAAGCAAGATTTGGGAAATAGAAATAAACCCAGGAGTGAATTTACTGGGTA-C 559

QY 419 AGCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAGAGGTGGAATACTTCAGATG 478  
DB 560 AGCAATTCAGCAACAGAGCTCTTCAGAAACTGAGGAGAGAGGTGGAATACTTCAGATG 619  
QY 479 CAACAGTGAAGAGAGAGGTGATGAGTGAAGAGATGGAAGAGGCAAGAAAGATG 538  
DB 620 CAACAGTGAAGAGAGAGGTGATGAGTGAAGAGATGGAAGAGGCAAGAAAGATG 676  
QY 539 AAAAGCAGGCTCAAAACGGAAGAGTCAATATCTTCAAGAAATCTCTTAAACAGTCCC 598  
DB 677 AAAAGCAGGCTCAAAACGGAAGAGTCTTCAACAGTCTTCTTAAACAGTCCC 736  
QY 599 GGAATCTCCAGGAGATGAAGATCAAGATCAAGAGTCTCAAGAGAGGAAACCAACAGCAGCT 658  
DB 737 GGAATCTCCAGGAGATGAAGATCAAGATCAAGAGTCTCAAGAGAGGAGAACCAACAGCAGCT 796  
QY 659 CTGAGGGTGGAGATCGGGGCAACGACACAGAAACACAACTTCAGACTTCGAGAAACCA 718  
DB 797 CTGAGGGGCGAGAGCGCTGGCAATGACACGAGAAACACGACTTCGAGAAAGCGG 856  
QY 719 GTGAGGGACCTAATACCATTAATGAATGCTGCAATTAAGAGAAACCAACAGAGGTTA 778  
DB 857 GTGAGGGACCTAATACCATTAATGAATGCTGCAATTAAGAGAAACCAACAGAGGTTA 916  
QY 779 TATGTTGGTGTCTAATATCTTGGATTTCATATGAACCAACACATAGTCTTCTGTTCTC 838  
DB 917 ACTGTTGATGCTAATCTTGGATTTCATATGAACCAACAGCTCTTCTGTTCTC 976  
QY 839 ATTGACAGAACCCAGTTTGTATGATTAATCATATTCCTCTCTCTGTTGTTGCGGG 898  
DB 977 ATTGACAGAGCCCGAGTATGTTGTCATTAAGTACATTCCTCTCTCTGTTGTTGGAAGA 1036  
QY 899 GGAAGAGCATTTAGCCTTTTAAAGTTACTGATTAATTTTTCATGTTTATTTGGTTGC 958  
DB 1037 AAAAGAGACATTTTAAACCTTTTAA-----CGTTGACTTTTATTTATTTGGTTGC 1091  
QY 959 ATGAAGTTGCTTAAACCACTAAGGATTATCAAGATTTTTCGCGAGACTTATACATGCT 1018  
DB 1092 ATGAAGTTGCTTAACTAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1148  
QY 1019 AGGATC 1024  
DB 1149 AGGCTC 1154

RESULT 10  
AK045542  
LOCUS  
DEFINITION  
AK045542 3363 bp mRNA linear HTC 05-DEC-2002  
Mus musculus adult male corpora quadrigemina cDNA, RIKEN  
full-length enriched library, clone:B230209J06  
product:hepatoma-derived growth factor, related protein 3, full  
insert sequence.  
ACCESSION  
VERSION AK045542.1 GI:26337454  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE  
1 Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
PUBLISHED 99279253  
10349636  
2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
JOURNAL  
MEDLINE

PUBMED  
REFERENCE  
AUTHORS

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, F., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Iehikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

20530913

4

11076861

4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Iehii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wegner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Togo-Oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

21085660

5

11217851

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

PUBMED  
REFERENCE  
AUTHORS

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, F., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Iehikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

20530913

4

11076861

4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Iehii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wegner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Togo-Oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

21085660

5

11217851

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

PUBMED  
REFERENCE  
AUTHORS

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, F., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Iehikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

20530913

4

11076861

4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Iehii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wegner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Togo-Oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

21085660

5

11217851

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

PUBMED  
REFERENCE  
AUTHORS

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, F., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Iehikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

20530913

4

11076861

4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Iehii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wegner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Togo-Oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

21085660

5

11217851

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

5

PUBMED  
REFERENCE  
AUTHORS

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yanamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, F., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Iehikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

20530913

4

11076861

4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Iehii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wegner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Togo-Oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

21085660

5

Db 913 GGAATCT-CAGGAGACGAGATGATAAAGACTGCAAGAGAGGAGAAACAAAGCAGCT 971  
QY 659 CTGAGGGTGGAGATCGGCAACGACACAGAAACACAACTTCAGACTTCGAGAAACCA 718  
Db 972 CTGAGGGCGGAGACGCTGGCAATGACACAGAAACACGACTGAGACTTCGCAAGACCG 1031  
QY 719 GTGAAGGACCTAACTACCAATGCTGATTAATGCTGATTAAGAGAAACCAACAGAGTTA 778  
Db 1032 GTGAAGGACCTAACTACCGTAATGCTGATTAAGAGAAACCAACAGAGTTA 1091  
QY 779 TATGTTGGTGTCTTAATATCTTGGATTGATTAAGAAACCAACATAGTCTTGTGTC 838  
Db 1092 ACTGTTGATGCTAACTTCTGGATTGATTAAGAAACCAACAGGCTTGTGTC 1151  
QY 839 ATTGACAGAACCCAGTTGATGATACATATTCATATTCCTCTGTTGTTGCTTCGGG 898  
Db 1152 ATTGACAGAACCCAGTTGATGATTAAGAAACCAACAGGCTTGTGTC 1211  
QY 899 GGAAGACATTTAGCTTTTAAAGTTACTGATTTAACTTCAATTTTCAATTTTGTGTC 958  
Db 1212 AAAGAGACATTTAACCCTTTAA-----GGTGAATTTATTTATTTATTTGTTGTC 1266  
QY 959 ATGAAGTTCCCTTAAACCACTAAGGATTAATCAAGATTTTGGCGAGACTTATACATGCT 1018  
Db 1267 ATGAAGTTCCCTTAAACCACTAAGGATTAATCAAGATTTTGGCGAGACTTATACATGCT 1323  
QY 1019 AGGATC 1024  
Db 1324 AGGCTC 1329

RESULT 11  
BG204817  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

819 bp mRNA linear EST 21-APR-2001  
BG204817 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.  
BG204817  
EST.  
BG204817.1 GI:13726504

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 819)  
Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,  
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J.,  
Lerner, L., Costanzo, D., McElligott, K., Booser, S., Mays, R., Smith,  
J., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K., Offenbacher,  
J., Danzig, J. and Ducar, M.

Creation of genome-wide protein expression libraries using random  
activation of gene expression

Nat. Biotechnol. 19 (5), 440-445 (2001)  
2127151  
11329013  
Contact: Scott J. Cain  
Athersys, Inc.  
3201 Carnegie Ave., Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scain@atersys.com

High quality sequence stop: 551.  
Location/Qualifiers  
1..819  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_lib="Athersys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression  
Libraries using Random Activation of Gene Expression',  
Nature Biotechnology, in press. Note that even though the  
cell type indicated is HT1080, since a random activation  
method was used, these sequence tags are not necessarily

FEATURES  
Source

RESULT 12  
BM722135  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS

692 bp mRNA linear EST 01-MAR-2002  
BM722135  
UI-E-E00-ahy-c-17-0-UI.r1 UI-E-E00 Homo sapiens cDNA clone  
UI-E-E00-ahy-c-17-0-UI 5', mRNA sequence.  
BM722135  
BM722135.1 GI:19042502  
EST.

BASE COUNT 289 a 149 c 180 g 199 t  
ORIGIN

expressed in HT1080 under normal circumstances."  
Query Match 72.2%; Score 739; DB 10; Length 819;  
Best Local Similarity 98.2%; Pred. No. 1.4e-122;  
Matches 778; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

QY 196 CCGGCCCGGATTGATGAACCTCCAGAGGGCGCTGTGAAGCCCTCCAGCAAAACAAGTATCT 255  
Db 31 CCGCTACAGATTGATGAACCTCCAGAGGGCGCTGTGAAGCCCTCCAGCAAAACAAGTATCT 90

QY 256 ATCTCTTTTGGCACCCCATGAACCTGCAATTTCTAGTCCCAAGACCTTTTCCATAT 315  
Db 91 ATCTCTTTTGGCACCCCATGAACCTGCAATTTCTAGTCCCAAGACCTTTTCCATAT 150

QY 316 AAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACAAACGGAAGGATTAAACGAAGATTG 375  
Db 151 AAGGAGTACAAAGACAAAGTTTGGAAAGTCAAAACAAACGGAAGGATTAAACGAAGATTG 209

QY 376 TGGGAAATAGAAAATAAACCCAGGAGTAAAGTTTACTGCTACCGGCAATTCAGCAACAG 435  
Db 210 TGGGAAATAGAAAATAAACCCAGGAGTAAAGTTTACTGCTACCGGCAATTCAGCAACAG 269

QY 436 AGCTCTTCAGAACTGAGGAGAGGTTGGAATATCTGCAGATGCAAGCAGTGAGGAAGAA 495  
Db 270 AGCTCTTCAGAACTGAGGAGAGGTTGGAATATCTGCAGATGCAAGCAGTGAGGAAGAA 329

QY 496 GGTGATAGATGAAGAAAGATGGAAGGCAAAAGAAAGATGAAAAGCAGGCTCAAAA 555  
Db 330 GGTGATAGATGAAGAAAGATGGAAGGCAAAAGAAAGATGAAAAGCAGGCTCAAAA 389

QY 556 CGGAAAAAGTTCATATCTTCAAGAAATCTCTTAAACAGTCCCGGAAATCTCCAGAGAT 615  
Db 330 CGGAAAAAGTTCATATCTTCAAGAAATCTCTTAAACAGTCCCGGAAATCTCCAGAGAT 449

QY 616 GAAGATGACAAAGACTGCAAAAGAGAGGAAACCAAAAGCAGCTCTGAGGGTGGAGATGCG 675  
Db 450 GAAGATGACAAAGACTGCAAAAGAGAGGAAACCAAAAGCAGCTCTGAGGGTGGAGATGCG 509

QY 676 GGCAACGACACAGAAACACAACTTCAGACTTGCAGAAACCAAGTGAAGGACCTAACTA 735  
Db 510 GGCAACGACACAGAAACACAACTTCAGACTTGCAGAAACCAAGTGAAGGACCTAACTA 569

QY 736 CCATTAATGAATGCTGCATATTAAGAGAAACCAACAGAGGTTATATGTTGTTGTTCTAA 795  
Db 570 CCATTAATGAATGCTGCATATTAAGAGAAACCAACAGAGGTTATATGTTGTTGTTCTAA 629

QY 796 TATTTCTGGATTGATGATGAACCAACACATAGTCTCTGTTGTCATTCAGAACCCCACT 855  
Db 630 TATTTCTGGATTGATGATGAACCAACACATAGTCTCTGTTGTCATTCAGAACCCCACT 689

QY 856 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 915  
Db 690 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 748

QY 916 CTTTCTTTTAAAGTTTACTGATTTAAATTTTCATGTTTTCATGTTGTTGTTGTTGTTTAC 975  
Db 749 C-TTTTCTTTTAAAGTTTACTGATTTAAATTTTCATGTTTTCATGTTGTTGTTGTTTAC 807

QY 976 CACTAAGGATTA 987  
Db 808 CACTAAGGATTA 819

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 692)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 8889548  
 COMMENT Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBR, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Dr. Gregg Hegeman  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 Seq primer: M13 Reverse.  
 FEATURES  
 source  
 1..692  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-800-ahy-c-17-0-UI"  
 /tissue\_type="fetal eye"  
 /dev\_stage="fetal"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-E-800"  
 /notes="Organ: eye; Vector: pT73-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 UI-E-800 is a cDNA library containing the following  
 tissue(s): fetal eye. The library was constructed  
 according to Bonaldo, Lennon and Soares, Genome Research,  
 6:791-806, 1996. First strand cDNA synthesis was primed  
 with an oligo-dT primer containing a Not I site. Double  
 stranded cDNA was ligated to an EcoR I adaptor, digested  
 with Not I, and cloned directionally into pT73-Pac  
 vector. The oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)18 tail. The  
 sequence tag for this library is CGGTATACC. This library  
 was created for the program, Gene Discovery in the Visual  
 System, supported by National Eye Institute (NEI)."  
 BASE COUNT 257 a 111 c 157 g 167 t  
 ORIGIN  
 Query Match 67.4%; Score 690.4; DB 12; Length 692;  
 Best Local Similarity 99.9%; Pred. No. 7.2e-114;  
 Matches 691; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 329 ACAAGTTTGCAAGTCAAAACGGAAGGATTTACGAGGATTTGGGAATAGAAA 388  
 DB 1 ACAAGTTTGCAAGTCAAAACGGAAGGATTTACGAGGATTTGGGAATAGAAA 60  
 QY 389 ATACCCAGGAGTAAAGTTTACTGGCTTACCGGCAATTCAGCAACAGAGCTCTTCAGAAA 448  
 DB 61 ATACCCAGGAGTAAAGTTTACTGGCTTACCGGCAATTCAGCAACAGAGCTCTTCAGAAA 120  
 QY 449 CTGAGGAGAGGTGGAAATCTCGAGATCAACAGCTGAGGAAGAGGTGATAGATAG 508  
 DB 121 CTGAGGAGAGGTGGAAATCTCGAGATCAACAGCTGAGGAAGAGGTGATAGATAG 180  
 QY 509 AAGAAGATGCAAAAGCGCAAAAGAAAGATGAAAGAGCAGGCTCAAAACGGAAGATCAT 568  
 DB 181 AAGAAGATGCAAAAGCGCAAAAGAAAGATGAAAGAGCAGGCTCAAAACGGAAGATCAT 240

QY 569 ATACTTTCAAAGAAATCCTCTAAACAGTCCCGGAAATCTCCAGGAGATGAAGATGACAAG 628  
 DB 241 ATACTTTCAAAGAAATCCTCTAAACAGTCCCGGAAATCTCCAGGAGATGAAGATGACAAG 300  
 QY 629 ACTGCAAGAGAGGAGGAAACAAAGAGAGCTCTGAGGGTGGAGATGCGGGCAACGACACAA 688  
 DB 301 ACTGCAAGAGAGGAGGAAACAAAGAGAGCTCTGAGGGTGGAGATGCGGGCAACGACACAA 360  
 QY 689 GAAACACAACTTCAGAGCTTGCAGAAAAACAGTGAAGGAGCTTAACCTACCAATAATGAATGC 748  
 DB 361 GAAACACAACTTCAGAGCTTGCAGAAAAACAGTGAAGGAGCTTAACCTACCAATAATGAATGC 420  
 QY 749 TGCATATTAAGAGAAACCAAGAGGTTATATGTTTGGTTGTCTAATAATTTCTTGGATTT 808  
 DB 421 TGCATATTAAGAGAAACCAAGAGGTTATATGTTTGGTTGTCTAATAATTTCTTGGATTT 480  
 QY 809 GATATGAACCAACACATAGTCTCTGTTGTCATTGACAGAACCCAGTTTGTATGTACATT 868  
 DB 481 GATATGAACCAACACATAGTCTCTGTTGTCATTGACAGAACCCAGTTTGTATGTACATT 540  
 QY 869 ATTCAATATTCCTCTCTGTTGTTGTTTCGGGGGAAAAAGACATTTTAGCCTTTTTTAAAAAGT 928  
 DB 541 ATTCAATATTCCTCTCTGTTGTTGTTTCGGGGGAAAAAGACATTTTAGCCTTTTTTAAAAAGT 600  
 QY 929 TACTGATTTAAATTTCAATGTTATTTGGTTGCATGAAGTTGCCCTTAACCACTAAGGATTAT 988  
 DB 601 TACTGATTTAAATTTCAATGTTATTTGGTTGCATGAAGTTGCCCTTAACCACTAAGGATTAT 660  
 QY 989 CAAGATTTTGGCGACACTTATACATGCTCTAG 1020  
 DB 661 CAAGATTTTGGCGACACTTATACATGCTCTAG 692

RESULT 13  
 BE886964  
 LOCUS  
 DEFINITION  
 601506447F1 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:3908171 5',  
 mRNA sequence.  
 ACCESSION BE886964  
 VERSION BE886964.1 GI:10341775  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 781)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM9719 row: n column: 12  
 High quality sequence stop: 766.  
 Location/Qualifiers  
 1..781  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3908171"  
 /issue\_type="leiomyosarcoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_71"  
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 2.1 kb."  
 BASE COUNT 285 a 129 c 168 g 199 t



FT Modified-site 149 /note= "Potential cAMP- and cGMP-dependent protein kinase phosphorylation site"

FT Modified-site 151 /note= "Potential protein kinase C phosphorylation site"

FT Modified-site 152 /note= "Potential protein kinase C phosphorylation site"

FT Modified-site 155 /note= "Potential protein kinase C phosphorylation site"

FT Modified-site 156 /note= "Potential cAMP- and cGMP-dependent protein kinase phosphorylation site"

FT Modified-site 159 /note= "Potential protein kinase C phosphorylation site"

FT Modified-site 162 /note= "Potential casein kinase II phosphorylation site"

FT Modified-site 176 /note= "Potential N-glycosylation site"

FT Modified-site 178 /note= "Potential casein kinase II phosphorylation site"

FT Modified-site 187 /note= "Potential N-glycosylation site"

FT Modified-site 191 /note= "Potential N-glycosylation site"

FT Modified-site 192 /note= "Potential casein kinase II phosphorylation site"

XX WO9938976-A2.

XX 05-AUG-1999.

XX 12-JAN-1999; 99WO-US000654.

XX 29-JAN-1998; 98US-0015412.

XX (INCY-) INCYTE PHARM INC.

XX Baughn M, Corley NC, Guegler KJ, Hillman JL, Lal P; Tang YT;

XX WPI; 1999-469331/39.

XX N-PSDB; AA200048.

XX New human growth factor homologs and their use in the diagnosis, treatment and prevention of cell proliferative and developmental disorders

XX Claim 1; Fig 2; 89pp; English.

XX This sequence is the HGFH2 Human Growth Factor Homologue 2. The HGFH2 protein has structural and chemical homology to the human hepatoma derived growth factor (HDPF). HGFH2 is 203 amino acids in length and has three potential N-glycosylation sites, two potential cAMP- and cGMP-dependent protein kinase phosphorylation sites, eight potential casein kinase II phosphorylation sites and six potential protein kinase C phosphorylation sites. HGFHs are expressed in cancerous tissues and appear to play a role in cell proliferative and developmental disorders. A purified antagonist of HGFH can be administered to a subject as a method of treating or preventing a cell proliferative disorder e.g. atherosclerosis, hepatitis, and some cancers. Antibodies against HGFH and antisense sequences may also be used as antagonist for treating the above. Developmental disorders treated by HGFH include renal tubular acidosis, anaemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, and epilepsy. HGFH antibodies and HGFH polynucleotides and polypeptides may also be used in various diagnostic methods.

XX SQ Sequence 203 AA;

Query Match 100.0%; Score 1086; DB 20; Length 203;

Best Local Similarity 100.0%; Pred. No. 2.5e-97;

Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARPRPREYKAGDLVFAKMGYPHWPARIDELPGAVKPPANKYPPIFFGTHETAFLGPK 60

DB 1 MARPRPREYKAGDLVFAKMGYPHWPARIDELPGAVKPPANKYPPIFFGTHETAFLGPK 60

QY 61 DLFPYKEYKDFGSKNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGEGNTADA 120

DB 61 DLFPYKEYKDFGSKNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGEGNTADA 120

QY 121 SSEEGRDVRVEDGKGRKNEKAGSKRKSYTSKKSKOSRKSPGDEDDKCKEENKSSS 180

DB 121 SSEEGRDVRVEDGKGRKNEKAGSKRKSYTSKKSKOSRKSPGDEDDKCKEENKSSS 180

QY 181 EGGDAGNDRNTTSDLOKTSEGT 203

DB 181 EGGDAGNDRNTTSDLOKTSEGT 203

RESULT 2

AAAB13521

ID AAB13521 standard; Protein; 203 AA.

XX AC AAB13521;

XX 13-NOV-2000 (first entry)

XX Human hepatoma-derived growth factor homologous polypeptide.

XX Human; hepatoma-derived growth factor homologous polypeptide; huHDPH; cell proliferation; haematopoiesis; angiogenesis; lymphocyte proliferation; infection; autoimmune disease; vascular disease; cancer.

XX Homo sapiens.

XX WO200037492-A2.

XX 29-JUN-2000.

XX 22-DEC-1999; 99WO-US30932.

XX 22-DEC-1998; 98US-0113344.

XX (ELIL ) LILLY & CO ELI.

XX Na S;

XX WPI; 2000-442638/38.

XX N-PSDB; AAA65004.

XX Novel human homologues of hepatoma-derived growth factor homologous nucleic acids, polypeptides useful as probes or amplification primers in the detection, quantitation or isolation of gene sequences or transcripts

XX Claim 9; Page 74-75; 75pp; English.

XX The present sequence is the human homologue of the hepatoma-derived growth factor (huHDPH). This protein is involved in tumour formation in some cells, and it is thought that the gene and protein will be useful in the diagnosis and treatment of infections, autoimmune disorders, vascular diseases and cancers. In addition, the gene can be used to produce transgenic animals which may be used as animal models for these diseases, and the protein can be used to induce cell proliferation, haematopoiesis, lymphocyte proliferation and angiogenesis.

XX SQ Sequence 203 AA;

Query Match 100.0%; Score 1086; DB 21; Length 203;

Best Local Similarity 100.0%; Pred. No. 2.5e-97;

Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARPRPREYKAGDLVFAKMGYPHWPARIDELPGAVKPPANKYPPIFFGTHETAFLGPK 60

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2004, 15:26:51 ; Search time 41 Seconds

(without alignments)  
785.890 Million cell updates/sec

Title: US-09-787-328B-4

Perfect score: 1086

Sequence: 1 MARPRREYKAGDLVFAKMK.....DAGNDTRNTTSLDQKTSST 203

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_19Jun03.\*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1086	100.0	203	AA228502	HGFH2 Human Growth
2	1086	100.0	203	AA213521	Human hepatoma-der
3	1086	100.0	203	AA288374	Human type II hepa
4	1086	100.0	203	AA282800	Human protein sequ
5	1086	100.0	203	AA270073	Human secreted pro
6	1086	100.0	203	AA265515	Human albumin fusi
7	1053.5	97.0	202	AA272327	Rat protein isolat
8	965	88.9	244	AA215333	Novel human diagno
9	940	86.6	203	AA213522	Human hepatoma-der

10	895	82.4	172	21	AA243542	Human cancer assoc
11	672.5	61.9	197	22	ABG15332	Novel human diagno
12	535.5	49.3	246	21	AA258173	Lung cancer associ
13	534.5	49.2	240	16	AA266727	Human hepatoma der
14	534.5	49.2	240	21	AA209984	Human HDGFIP prote
15	534.5	49.2	240	22	AB250263	Hepatoma-derived g
16	532.5	49.0	235	19	AA261963	Lung growth factor
17	532	49.0	237	18	AA233216	Mouse hepatoma der
18	525	48.3	326	22	AB212116	Human PRO1604 homo
19	524	48.3	314	22	AA270075	Human secreted pro
20	524	48.3	314	23	AB265517	Human albumin fusi
21	524	48.3	671	21	AA299426	Human PRO1604 (UNQ
22	524	48.3	671	22	AA291996	Human PRO polypept
23	524	48.3	671	22	AA266175	Protein of the inv
24	524	48.3	671	23	AA297038	Human LP protein I
25	524	48.3	671	23	AB250542	Human NOV4d protei
26	524	48.3	671	24	AB271284	Human PRO1604 prot
27	524	48.3	671	24	AB265741	Human secreted/tra
28	524	48.3	671	24	AB266074	Novel human secret
29	524	48.3	671	24	AB267578	Human secreted/tra
30	524	48.3	671	24	AB265436	Human PRO polypept
31	524	48.3	671	24	AB258572	Human PRO polypept
32	524	48.3	671	24	AB256108	Human secreted/tra
33	524	48.3	671	24	AB257103	Human PRO polypept
34	524	48.3	671	24	AB210682	Human secreted/tra
35	524	48.3	676	22	AA264373	Human hepatome cel
36	523	48.2	676	22	AA247218	Human NOV9 protein
37	522	48.1	667	23	AB250540	Human NOV4b protei
38	522	48.1	667	23	AB250541	Human NOV4c protei
39	504.5	46.5	669	19	AA237483	Mouse liver cancer
40	504	46.4	676	22	AA209441	Human 'SB-HDGF prot
41	503	46.3	676	23	AB250539	Human NOV4a protei
42	486.5	44.8	530	20	AA297775	Human lens epithel
43	486.5	44.8	530	23	AA274502	Human lens epithel
44	485.5	44.7	325	21	AA294463	Human p75 protein
45	485.5	44.7	333	21	AA294462	Human p52 protein.

#### ALIGNMENTS

RESULT 1

AA228502

ID AA228502 standard; Protein; 203 AA.

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;

XX AC AA228502;



Db 1 MARPPREYKAGDLVFAKMGYPHWPARIIDELPEGAVKPPANKYPIFFGTHETAFLGPK 60  
 QY 61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGGGNTADA 120  
 Db 61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGGGNTADA 120  
 QY 121 SSEEGRVVEDGKGRKNEKAGSKRKYSYTSKKSSQSKRSKSPGDEDDKCKEENKSSS 180  
 Db 121 SSEEGRVVEDGKGRKNEKAGSKRKYSYTSKKSSQSKRSKSPGDEDDKCKEENKSSS 180  
 QY 181 EGGDAGNDRTRNTTSDLQKTSEGT 203  
 Db 181 EGGDAGNDRTRNTTSDLQKTSEGT 203

RESULT 3  
 AAY88374  
 ID AAY88374 standard; Protein; 203 AA.  
 AC AAY88374;  
 XX  
 DT 20-JUL-2000 (first entry)  
 XX  
 DE Human type II hepatoma derived growth factor (HDGF2) amino acid sequence.  
 XX  
 KW Hepatoma-derived growth factor 2; HDGF2; cytostatic; drug;  
 XX treatment; hepatoma.  
 OS Homo sapiens.  
 XX  
 PN W0200017351-A1.  
 XX  
 PD 30-MAR-2000.  
 XX  
 PF 06-SEP-1999; 99WO-CN00139.  
 XX  
 PR 22-SEP-1998; 98CN-0119758.  
 XX  
 PA (YULL/) YU L.  
 XX  
 PI Yu L, Zhang H, Fu Q, Zhao Y, Tu Q;  
 XX WPI; 2000-283579/24.  
 DR N-PSDB; AAA13161.  
 XX

Type II human-derived growth factor with homology to its type I version, useful e.g. in study and development of drugs for hepatoma -  
 PS Claim 2; Page 15; 27pp; Chinese.  
 XX  
 CC This sequence represents the human hepatoma-derived growth factor (HDGF2) protein sequence. The HDGF2 sequence shows considerable homology to the type I version. The protein has cytostatic activity. The HDGF2 protein, derivatives and polynucleotides are useful e.g. in the study and development of drugs for the treatment of hepatoma.  
 XX  
 SQ Sequence 203 AA;

Query Match 100.0%; Score 1086; DB 21; Length 203;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-97;  
 Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARPPREYKAGDLVFAKMGYPHWPARIIDELPEGAVKPPANKYPIFFGTHETAFLGPK 60  
 Db 1 MARPPREYKAGDLVFAKMGYPHWPARIIDELPEGAVKPPANKYPIFFGTHETAFLGPK 60  
 QY 61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGGGNTADA 120  
 Db 61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGGGNTADA 120  
 QY 121 SSEEGRVVEDGKGRKNEKAGSKRKYSYTSKKSSQSKRSKSPGDEDDKCKEENKSSS 180  
 Db 121 SSEEGRVVEDGKGRKNEKAGSKRKYSYTSKKSSQSKRSKSPGDEDDKCKEENKSSS 180

QY 181 EGGDAGNDRTRNTTSDLQKTSEGT 203  
 Db 181 EGGDAGNDRTRNTTSDLQKTSEGT 203

RESULT 4  
 AAB92800  
 ID AAB92800 standard; Protein; 203 AA.  
 AC AAB92800;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human protein sequence SEQ ID NO:11307.  
 XX  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 XX  
 PR 27-AUG-1999; 99JP-0300253.  
 XX  
 PR 11-JAN-2000; 2000JP-0118776.  
 XX  
 PR 02-MAY-2000; 2000JP-0183767.  
 XX  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX WPI; 2001-318749/34.  
 DR  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -  
 PT  
 XX  
 PS Claim 8; SEQ ID 11307; 2537pp + CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AA03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.  
 XX  
 SQ Sequence 203 AA;

Query Match 100.0%; Score 1086; DB 22; Length 203;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-97;

Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARPPREYKAGDLVFAKMGYPHPWPARIDELPEGAVKPPANKYPPIFFGTHETAFLGPK 60  
DB 1 MARPPREYKAGDLVFAKMGYPHPWPARIDELPEGAVKPPANKYPPIFFGTHETAFLGPK 60

QY 61 DLFPYKEYDKFGKSNKRKGFNEGLWEIENNPVGFYQAIQQOSSSETEGEGNTADA 120  
DB 61 DLFPYKEYDKFGKSNKRKGFNEGLWEIENNPVGFYQAIQQOSSSETEGEGNTADA 120

QY 121 SSEEGRVVEEDGKGRKNEKAGSKRKSYTSKSSKQSRKSPGDEDDKCKEENKSSS 180  
DB 121 SSEEGRVVEEDGKGRKNEKAGSKRKSYTSKSSKQSRKSPGDEDDKCKEENKSSS 180

QY 181 EGGDAGNDRNTTSDLOKTSEGT 203  
DB 181 EGGDAGNDRNTTSDLOKTSEGT 203

RESULT 5  
AAB70073  
ID AAB70073 standard; Protein; 203 AA.  
AC AAB70073;  
DT 14-MAY-2001 (first entry)  
XX Human secreted protein #12.  
DE Human; secreted protein; immunomodulatory; antisclerotic;  
KW dermatological; immunosuppressive; antiinflammatory; anti-HIV;  
KW immunostimulant; cytostatic; cardiant; vascular; anti-angiogenic;  
KW ophthalmological; neuroprotectant; nootropic; anticonvulsant; vaccine;  
KW antialzheimers; antiparkinsonian; antimicrobial; vulnary; gene therapy;  
KW immune disorder; hyperproliferative; cardiovascular; angiogenic;  
KW neurological; infection.  
XX Homo sapiens.  
XX WO200112776-A2.  
XX 22-FEB-2001.  
XX 15-AUG-2000; 2000WO-US22350.  
XX 16-AUG-1999; 99US-0148759.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Shi Y, Young PE, Ebner R, Soppet DR, Ruben SM;  
XX WPI; 2001-244245/25.  
XX N-PSDB; AAF76854.  
XX Nucleic acids encoding 18 human secreted polypeptides, useful for  
XX preventing, diagnosing and/or treating e.g. cancers, Parkinson's  
XX disease and diabetic retinopathy -  
XX Claim 11; Page 369; 380pp; English.

The present sequence is one of 18 novel human secreted proteins. The  
nucleic acids encoding the proteins and the proteins themselves may be  
used in the prevention, diagnosis and treatment of diseases including  
immune disorders (e.g. multiple sclerosis, systemic lupus erythematosus  
and human immunodeficiency virus (HIV) infections), hyperproliferative  
disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases  
(e.g. Scimitar syndrome, Chaga's cardiomyopathy and coronary  
arteriosclerosis), angiogenic disorders (e.g. corneal graft  
neovascularisation and diabetic retinopathy), neurological disorders  
(e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),  
infectious diseases and/or for promoting wound healing, regeneration  
and/or chemotaxis. The nucleic acid molecules may also be used as DNA  
probes in diagnostic assays to detect and quantitate the presence of

CC similar nucleic acid sequences in samples. The polypeptides may also be  
used as antigens in the production of antibodies and in assays to  
identify modulators of protein expression and activity.

XX Sequence 203 AA;

Query Match 100.0%; Score 1086; DB 22; Length 203;  
Best Local Similarity 100.0%; Pred. No. 2.5e-97;  
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARPPREYKAGDLVFAKMGYPHPWPARIDELPEGAVKPPANKYPPIFFGTHETAFLGPK 60  
DB 1 MARPPREYKAGDLVFAKMGYPHPWPARIDELPEGAVKPPANKYPPIFFGTHETAFLGPK 60

QY 61 DLFPYKEYDKFGKSNKRKGFNEGLWEIENNPVGFYQAIQQOSSSETEGEGNTADA 120  
DB 61 DLFPYKEYDKFGKSNKRKGFNEGLWEIENNPVGFYQAIQQOSSSETEGEGNTADA 120

QY 121 SSEEGRVVEEDGKGRKNEKAGSKRKSYTSKSSKQSRKSPGDEDDKCKEENKSSS 180  
DB 121 SSEEGRVVEEDGKGRKNEKAGSKRKSYTSKSSKQSRKSPGDEDDKCKEENKSSS 180

QY 181 EGGDAGNDRNTTSDLOKTSEGT 203  
DB 181 EGGDAGNDRNTTSDLOKTSEGT 203

RESULT 6  
ABG65515  
ID ABG65515 standard; Protein; 203 AA.  
AC ABG65515;  
XX 27-AUG-2002 (first entry)  
DT Human albumin fusion protein #2190.  
DE Human albumin fusion protein #2190.  
XX Albumin fusion protein; therapeutic protein X; human albumin; HA;  
KW human serum albumin; HSA; cancer; reproductive disorder;  
KW digestive disorder; immune disorder; endocrine disorder;  
KW haematopoietic disorder; neural disorder; connective disorder;  
KW cytostatic; antiinfertility; antiinflammatory; antitumor;  
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;  
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;  
KW osteopathic; antiarthritic.  
XX Homo sapiens.  
XX Synthetic.  
XX WO200177137-A1.  
XX PD 18-OCT-2001.  
XX 12-APR-2001; 2001WO-US11988.  
XX 12-APR-2000; 2000US-229358P.  
XX 25-APR-2000; 2000US-199384P.  
XX 21-DEC-2000; 2000US-256931P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Haseltine WA;  
XX WPI; 2002-010886/01.  
XX New fusion protein for treating disease e.g. diabetes comprises an  
XX albumin fused to a therapeutic protein -  
XX Claim 1; Page 2083-2084; 2102pp; English.  
XX The present invention relates to albumin fusion proteins comprising a  
XX therapeutic protein X and human albumin (HA), also known as human serum  
XX albumin, HSA). The proteins are useful for treating a disease or

CC disorder that may be modulated by therapeutic protein X. The albumin  
 CC extends the shelf-life of protein X, and may increase its biological  
 CC in vitro/in vivo activity. The protein is useful for treating and  
 CC diagnosing disorders such as cancer, reproductive disorders, digestive  
 CC disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders  
 CC (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders  
 CC (e.g. diabetes), haematopoietic disorders, neural disorders  
 CC (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,  
 CC encephalomyelitis, meningitis, schizophrenia), and connective disorders  
 CC (e.g. osteoporosis, arthritis). ABG63326-ABG65518 represent albumin  
 CC fusion proteins of the invention.  
 XX  
 SQ Sequence 203 AA;

Query Match 100.0%; Score 1086; DB 23; Length 203;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-97;  
 Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MARPPREYKAGDLVFAKMGYPHPARIDELPEGAVKPPANKYPIFFFGTHETAFILGPK 60  
 DB 1 MARPPREYKAGDLVFAKMGYPHPARIDELPEGAVKPPANKYPIFFFGTHETAFILGPK 60  
 QY 61 DLPPYKEYKDKFGKSKRKGFGNEGLWEIENNPVGKFTGYQAIQQSSSETEGEGNTADA 120  
 DB 61 DLPPYKEYKDKFGKSKRKGFGNEGLWEIENNPVGKFTGYQAIQQSSSETEGEGNTADA 120  
 QY 121 SSEEGRVVEDGKGRKNEKAGSKRKSYTSKSKSKSRKSPGDEDDKCKEENKSSS 180  
 DB 121 SSEEGRVVEDGKGRKNEKAGSKRKSYTSKSKSKSRKSPGDEDDKCKEENKSSS 180  
 QY 181 EGGDAGNDTRNTTSDLQKTSEGT 203  
 DB 181 EGGDAGNDTRNTTSDLQKTSEGT 203

RESULT 7  
 ABB72327  
 ID ABB72327 standard; Protein; 202 AA.  
 XX  
 AC ABB72327;  
 XX  
 DT 04-APR-2002 (first entry)  
 XX  
 DE Rat protein isolated from skin cells SEQ ID NO: 651.  
 XX  
 KW Human; rat; mouse; skin call; skin wound; cancer; growth defect;  
 KW developmental defect; inflammatory disease; dermatological; vulnary;  
 KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.  
 XX  
 OS Rattus sp.  
 XX  
 PN WO200190357-A1.  
 XX  
 PD 29-NOV-2001.  
 XX  
 PF 24-MAY-2001; 2001WO-NZ00099.  
 XX  
 PR 24-MAY-2000; 2000US-206650P.  
 PR 25-JUL-2000; 2000US-221232P.  
 XX  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 XX  
 PI Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG, Kumble KD;  
 PI WPI; 2002-122020/16.  
 DR N-PSDB; ABLJ5012.  
 DR  
 XX New polynucleotides and polypeptides encoded by the polynucleotides  
 PT isolated from skin cells, useful for treating skin wounds, cancers,  
 PT growth and developmental defects, inflammatory diseases, or for  
 PT modulating immune responses -  
 XX  
 PS Claim 4; Page 408-409; 466pp; English.

XX The present invention provides the protein and coding sequences of cDNAs  
 CC isolated from human, murine and rat skin cell libraries. The sequences  
 CC can be used in the development of therapeutic agents useful in the  
 CC treatment of skin diseases, including skin wounds, cancer, growth  
 CC defects, developmental defects and inflammatory diseases. The proteins  
 CC have important roles in the induction of hair growth, cell proliferation  
 CC and cell-cell interaction, in maintaining tissue integrity, in wound  
 CC healing and in modulating immune responses. The present sequence is a  
 CC polypeptide of the invention.  
 XX  
 SQ Sequence 202 AA;

Query Match 97.0%; Score 1053.5; DB 23; Length 202;  
 Best Local Similarity 97.5%; Pred. No. 3.5e-94;  
 Matches 198; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
 QY 1 MARPPREYKAGDLVFAKMGYPHPARIDELPEGAVKPPANKYPIFFFGTHETAFILGPK 60  
 DB 1 MARPPREYKAGDLVFAKMGYPHPARIDELPEGAVKPPANKYPIFFFGTHETAFILGPK 60  
 QY 61 DLPPYKEYKDKFGKSKRKGFGNEGLWEIENNPVGKFTGYQAIQQSSSETEGEGNTADA 120  
 DB 61 DLPPYKEYKDKFGKSKRKGFGNEGLWEIENNPVGKFTGYQAIQQSSSETEGEGNTADA 120  
 QY 121 SSEEGRVVEDGKGRKNEKAGSKRKSYTSKSKSKSRKSPGDEDDKCKEENKSSS 180  
 DB 121 SSEEGRVVEDGKGRKNEKAGSKRKSYTSKSKSKSRKSPGDEDDKCKEENKSSS 180  
 QY 181 EGGDAGNDTRNTTSDLQKTSEGT 203  
 DB 180 EGGDAGNDTRNTTSDLQKAGEGT 202

RESULT 8  
 ABG15333  
 ID ABG15333 standard; Protein; 244 AA.  
 XX  
 AC ABG15333;  
 XX  
 DT 18-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #15324.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US08631.  
 XX  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 PI WPI; 2001-639362/73.  
 DR N-PSDB; AAS79520.  
 DR  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -  
 XX  
 PS Claim 20; SEQ ID No 45692; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and



CC involved in tumour formation in some cells, and it is thought that the  
 CC gene and protein will be useful in the diagnosis and treatment of  
 CC infections, autoimmune disorders, vascular diseases and cancers. In  
 CC addition, the gene can be used to produce transgenic animals which may be  
 CC used as animal models for these diseases, and the protein can be used to  
 CC induce cell proliferation, haematopoiesis, lymphocyte proliferation and  
 CC angiogenesis.  
 CC Note: This sequence is not shown in the specification, but is derived  
 CC from the wild-type human homologue of the hepatoma-derived growth factor  
 CC shown in SEQ ID NO: 2 (AAB13521).

XX  
 SQ Sequence 203 AA;  
 Query Match 86.6%; Score 940; DB 21; Length 203;  
 Best Local Similarity 85.2%; Pred. No. 3.7e-83;  
 Matches 173; Conservative 11; Mismatches 19; Indels 0; Gaps 0;  
 QY 1 MARPRPKAGDLVFAKMGKYPHPARIDLPAGAVKPPANKYPIFFFGTHETAFILGPK 60  
 DB 1 MARPRPREFKPGDLIFAKMGKYPHPARVDDXADGAVKPPXNKLPIFFFGTHETAFILGPK 60  
 QY 61 DLPYKPKYKDFKGSNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGEGNTADA 120  
 DB 61 DLPYKPKYKDFKGSNKRKGFNEGLWEIENNPXASYSXXPPIXSSSETEGEGNTADA 120  
 QY 121 SSEEGRVREDGKGRKNEKAGSKRKYTSKSSKQSRKSPGDEDDKCKEENKSSS 180  
 DB 121 SSEEGRVREDGKGRKNEKAGSKRKYTSKSSKQSRKSPGDEDDKCKEENKSSS 180  
 QY 181 EGGDAGNDTRNTTSDLOKTSSEGT 203  
 DB 181 EGGDAGNDTRNTTSDLOKTSSEGT 203

RESULT 10  
 AAB43542  
 ID AAB43542 standard; Protein; 172 AA.  
 XX  
 AC AAB43542;  
 XX  
 DT 08-FEB-2001 (first entry)  
 XX  
 DE Human cancer associated protein sequence SEQ ID NO:987.  
 XX  
 KW Human; cancer associated gene; cancer antigen; detection; cancer;  
 KW diagnosis; cytostatic; proliferative; vulnary; immunomodulator;  
 KW antidiabetic; antisthmatic; antirheumatic; antithyroid; antitumor; antiviral;  
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;  
 KW dermatological; neuroprotective; thrombolytic; coagulant; neutropic;  
 KW vasotrophic; antipsoriatic; antiangiogenic; gene therapy; inflammation;  
 KW immune disorder; haematopoietic cell disorder; autolimmune disorder;  
 KW allergic reaction; graft versus host disease; organ rejection;  
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;  
 KW neurological disease; drug screening.  
 OS  
 XX Homo sapiens.  
 XX OS  
 XX WO200005350-A1.  
 XX  
 XX PD 21-SEP-2000.  
 XX  
 XX PF 08-MAR-2000; 2000WO-US05882.  
 XX  
 XX PR 12-MAR-1999; 99US-0124270.  
 XX  
 XX PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX PI Rosen CA, Ruben SM;  
 XX  
 XX DR WPI; 2000-587533/55.  
 XX  
 XX DR N-PSDB; AAC77751.  
 XX  
 XX PT Novel isolated nucleic acids comprising sequences encoding peptides

PT useful for treating or diagnosing e.g. cancer -  
 XX Claim 11; Page 1562; 2352pp; English.  
 XX  
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given  
 CC in AAB43398 to AAB44239. The proteins can have activities based on the  
 CC tissues and cells the genes are expressed in. Example of activities  
 CC include: cytostatic; proliferative; vulnary; immunomodulator;  
 CC antidiabetic; antisthmatic; antirheumatic; antithyroid; antitumor;  
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;  
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;  
 CC neutropic; vasotropic; antipsoriatic and antiangiogenic. The  
 CC polynucleotides and polypeptides can be used for preventing, treating or  
 CC ameliorating medical conditions and diagnosing pathological conditions.  
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from  
 CC the present invention may be used to treat immune disorders by activating  
 CC or inhibiting the proliferation, differentiation or mobilisation of  
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune  
 CC disorders, allergic reactions, graft versus host disease and organ  
 CC rejection, modulate haemostatic or thrombolytic activity, modulate  
 CC inflammation, cancers, cardiovascular disorders, neurological disease and  
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,  
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to  
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of  
 CC the present invention.

XX Sequence 172 AA;  
 SQ  
 Query Match 82.4%; Score 895; DB 21; Length 172;  
 Best Local Similarity 98.8%; Pred. No. 6.9e-79;  
 Matches 169; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 33 PEGAVKPPANKYPIFFFGTHETAFILGPKDLPYKFTGYQAIQQSSSETEGEGNTADA 152  
 DB 2 PRGAVKPSANKYPIFFFGTHETAFILGPKDLPYKFTGYQAIQQSSSETEGEGNTADA 152  
 QY 93 GVKFTGYQAIQQSSSETEGEGNTADASSEEGRVREDGKGRKNEKAGSKRKYTS 152  
 DB 62 GVKFTGYQAIQQSSSETEGEGNTADASSEEGRVREDGKGRKNEKAGSKRKYTS 121  
 QY 153 KSSSKSRKSPGDEDDKCKEENKSSSEGGDAGNDTRNTTSDLOKTSSEGT 203  
 DB 122 KSSSKSRKSPGDEDDKCKEENKSSSEGGDAGNDTRNTTSDLOKTSSEGT 172

RESULT 11  
 ABG15332  
 ID ABG15332 standard; Protein; 197 AA.  
 XX  
 AC ABG15332;  
 XX  
 DT 18-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #15323.  
 XX  
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200175067-A2.  
 XX  
 XX PD 11-OCT-2001.  
 XX  
 XX PF 30-MAR-2001; 2001WO-US08631.  
 XX  
 XX PR 31-MAR-2000; 2000US-0540217.  
 XX  
 XX PR 23-AUG-2000; 2000US-0649167.  
 XX  
 XX PA (HYSE-) HYSEQ INC.  
 XX  
 XX PI Drmanac RT, Liu C, Tang YT;  
 XX

DR WPI; 2001-639362/73.  
 DR N-PSDB; AAS79519.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -  
 XX  
 PS Claim 20; SEQ ID No 45691; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AEG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: the sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 197 AA;

Query Match 61.9%; Score 672.5; DB 22; Length 197;  
 Best Local Similarity 87.2%; Pred. No. 3.3e-57;  
 Matches 129; Conservative 3; Mismatches 7; Indels 9; Gaps 1;

QY 19 MKGYHPHWPAR-----IDELPGAVKPPANKYPIFFFGTHETAFGLGPKDLFPYKVK 69  
 Db 35 LRAAPLQPTRDPSDMVAFIDELPGAVKPPANKYPIFFFGTHETAFGLGPKDLFPYKVK 94  
 QY 70 DFGKSNKRGKGFNEGLWEIENNPVGKFTGYQAIQQSSSTETEGEGNTADASSEEGDRV 129  
 Db 95 DFGKSNKRGKGFNEGLWEIENNPVGKFTGYQAIQQSSSTETEGEGNTADASSEEGDRV 154  
 QY 130 EDGKGRKNEKAGSKRKKSYTSKSSK 157  
 Db 155 EDGKGRKNEKAGSKRKKSYTSKTK 182

RESULT 12  
 AAB58173  
 ID AAB58173 standard; Protein; 246 AA.  
 XX  
 AC AAB58173;  
 XX  
 DT 14-MAR-2001 (first entry)  
 XX  
 DE Lung cancer associated polypeptide sequence SEQ ID 511.  
 DE  
 KW Human; lung cancer associated protein; neuroprotective; cytostatic;  
 KW cardioactive; immunomodulatory; muscular active; vulnerary;  
 KW gastrointestinal; nephrotropic; antiinfective; gynecological;  
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
 KW proliferative disorder; wound healing; infectious disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200055180-A2.  
 XX  
 PD 21-SEP-2000.  
 XX

PF 08-MAR-2000; 2000WO-US05918.  
 XX  
 PR 12-MAR-1999; 99US-0124270.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (ROSE/) ROSEN C A.  
 XX  
 PI Ruben SM;  
 XX  
 DR WPI; 2000-587514/55.  
 DR N-PSDB; AAF18049.  
 XX  
 PT Lung cancer associated gene sequences, referred to as lung cancer  
 PT antigens, useful for treatment, prevention, and diagnosis of disorders  
 PT such as lung cancer -  
 XX  
 PS Claim 11; Page 999-1000; 1425pp; English.  
 XX  
 CC Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer  
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
 CC associated proteins and polynucleotide sequences, their agonists, and  
 CC antagonists may have neuroprotective; cytostatic; cardioactive;  
 CC immunomodulatory; muscular active general; vulnerary; gastrointestinal  
 CC general; nephrotropic; antiinfective; gynecological; or antibacterial  
 CC activity. The invention also includes antibodies specific for the  
 CC protein or polynucleotide sequences. The lung cancer associated  
 CC polynucleotide sequences may be used for detection of lung cancer,  
 CC chromosome identification, as chromosome markers, and for numerous other  
 CC diagnostic or research purposes. The proteins may be used to treat  
 CC disorders such as neural, immune, muscular, reproductive,  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders. The proteins may also be used in the treatment of wounds and  
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and  
 CC peptide AAB58549 are used in the course of the invention for the  
 CC identification and characterisation of the polynucleotide and protein  
 CC sequences.  
 XX  
 SQ Sequence 246 AA;

Query Match 49.3%; Score 535.5; DB 21; Length 246;  
 Best Local Similarity 51.4%; Pred. No. 8.7e-44;  
 Matches 114; Conservative 25; Mismatches 58; Indels 25; Gaps 6;

QY 5 RPREYKAGDLVFAKMKGYHPHWPARIDELPGAVKPPANKYPIFFFGTHETAFGLGPKDLFP 64  
 Db 12 RQKEYKCGDLVFAKMKGYHPHWPARIDENPEAAVKSTANKYQVFFGTHETAFGLGPKDLFP 71  
 QY 65 YKEYKDKFGKSKRGKGFNEGLWEIENNPVGKFTGYQAIQQSSSSE-----TEGEG-- 114  
 Db 72 YEESEKFKGKPKRKGFGSEGLWEIENNPVTKASGYQSSQKSKSCVPEPEPEAAEGDGDK 131  
 QY 115 -GNTADASSEEGDRVVEEDGKGRKNEKAGSKRKKSKS---YTSKSKSKSRKSPGDEDDKD 170  
 Db 132 KGN-AEGSSDBEG-KLVIDEPAKNEKNGKALGRAGDLDLEDSPKPKPAENPEGEKEAA 189  
 QY 171 CKE-----EENKSSSEGGDAGNDRNTTSDLOKTTSEGT 203  
 Db 190 TLEVERPLPMEVEKNSTXSEFGSGRGPQEESEEEDEEEAT 231

RESULT 13  
 AAR66727  
 ID AAR66727 standard; Protein; 240 AA.  
 XX  
 AC AAR66727;  
 XX  
 DT 02-OCT-1995 (first entry)  
 XX  
 DE Human hepatoma derived growth factor.  
 XX  
 KW Human hepatoma derived growth factor; recombinant production;  
 KW receptor purification.  
 XX

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2004, 15:26:53 ; Search time 22 seconds  
(without alignments)  
390.414 Million cell updates/sec

Title: US-09-787-328B-4

Perfect score: 1086

Sequence: 1 MARPPREYKAGDLVFAKMK.....DAGNDRNTTDLQKTSEGT 203

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A COMB.pcp.\*

2: /cgn2\_6/ptodata/1/iaa/5B COMB.pcp.\*

3: /cgn2\_6/ptodata/1/iaa/6A COMB.pcp.\*

4: /cgn2\_6/ptodata/1/iaa/6B COMB.pcp.\*

5: /cgn2\_6/ptodata/1/iaa/PTUS COMB.pcp.\*

6: /cgn2\_6/ptodata/1/iaa/backfile1.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	534.5	49.2	240	US-08-760-745-3	Sequence 3, Appli
2	532.5	49.0	235	US-08-760-745-1	Sequence 1, Appli
3	532	49.0	237	US-08-760-745-5	Sequence 5, Appli
4	133.5	12.3	764	US-09-370-838-67	Sequence 67, Appl
5	119.5	11.0	302	US-09-282-305-14	Sequence 14, Appl
6	119.5	11.0	302	US-09-883-720-14	Sequence 14, Appl
7	118	10.9	305	US-09-282-305-12	Sequence 12, Appl
8	118	10.9	305	US-09-883-720-12	Sequence 12, Appl
9	117.5	10.8	414	US-07-667-276A-4	Sequence 4, Appli
10	114	10.5	445	US-08-845-258-38	Sequence 38, Appl
11	114	10.5	445	US-08-950-571-38	Sequence 38, Appl
12	114	10.5	445	US-08-723-142A-38	Sequence 38, Appl
13	114	10.5	445	US-09-528-784A-38	Sequence 38, Appl
14	114	10.5	445	US-09-569-098A-38	Sequence 38, Appl
15	114	10.5	666	US-09-528-784A-85	Sequence 85, Appl
16	114	10.5	666	US-09-569-098A-85	Sequence 85, Appl
17	114	10.5	1132	US-09-528-784A-87	Sequence 87, Appl
18	114	10.5	1132	US-09-569-098A-87	Sequence 87, Appl
19	112	10.3	231	US-09-461-697-194	Sequence 194, App
20	112	10.3	232	US-09-461-697-192	Sequence 192, App
21	112	10.3	238	US-09-461-697-190	Sequence 190, App
22	112	10.3	257	US-09-461-697-188	Sequence 188, App
23	112	10.3	272	US-09-461-697-186	Sequence 186, App
24	112	10.3	311	US-09-282-305-16	Sequence 16, Appl
25	112	10.3	311	US-09-883-720-16	Sequence 16, Appl
26	111.5	10.3	482	US-08-528-255A-1	Sequence 1, Appli
27	111.5	10.3	482	US-08-717-365-1	Sequence 1, Appli

28	109.5	10.1	415	3	US-08-675-816-2	Sequence 2, Appli
29	107.5	9.9	1085	1	US-08-431-080-28	Sequence 28, Appl
30	107.5	9.9	1085	2	US-08-938-534-28	Sequence 28, Appl
31	107.5	9.9	1085	4	US-09-345-294-28	Sequence 28, Appl
32	105.5	9.7	1850	4	US-09-480-921B-8	Sequence 8, Appli
33	105.5	9.7	1850	4	US-09-620-093A-5	Sequence 5, Appli
34	102.5	9.4	913	4	US-08-971-089-4	Sequence 4, Appli
35	100	9.2	531	2	US-08-933-750C-9	Sequence 9, Appli
36	100	9.2	531	3	US-09-234-613-9	Sequence 9, Appli
37	99	9.1	210	4	US-09-252-991A-19969	Sequence 19969, A
38	98	9.0	317	6	5340934-11	Patent No. 5340934
39	98	9.0	754	3	US-09-214-564A-2	Sequence 2, Appli
40	97.5	9.0	1162	2	US-08-728-323A-2	Sequence 2, Appli
41	97.5	9.0	1162	4	US-09-298-568-2	Sequence 2, Appli
42	97.5	9.0	1162	4	US-09-410-399-2	Sequence 2, Appli
43	97	8.9	182	1	US-08-466-603-5	Sequence 5, Appli
44	97	8.9	182	1	US-08-314-503A-5	Sequence 5, Appli
45	97	8.9	182	1	US-08-468-066-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1  
US-08-760-745-3  
; Sequence 3, Application US/08760745  
; Patent No. 5972658  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Goli, Surya K.  
; APPLICANT: Murty, Lynn E.  
; TITLE OF INVENTION: LUNG GROWTH FACTOR VARIANT  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/760,745  
; FILING DATE: Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0169 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 240 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 598956  
; US-08-760-745-3  
Query Match 49.2%; Score 534.5; DB 2; Length 240;  
Best Local Similarity 51.4%; Pred. No. 1.6e-46;

Matches	114;	Conservative	25;	Mismatches	58;	Indels	25;	Gaps	67;
5	RPREYKAGDLVFAKMKGYPHWP	PARIDELPEGA	VPPANKYPIFF	FGTGHETAFLGPKDLFP	64				
6	RQKEYKGDVLFVAKMKGYPHWP	PARIDEMPEAA	VKSTANKYQVFF	FGTGHETAFLGPKDLFP	65				
65	YKEYDKKGGKNKRKGFNEGLWE	TENNPGVFTGYQAI	QQQSSSE	-----TEGEG--	114				
66	YEESEKFGKGNKRKGFSEGLWE	TENNPTVKASGY	QQSQKKSCVEE	PEPPEAAEGGDGX	125				
115	-GNTADASSEEGDVRVEDGK	RKVEKAGSKRKKS	---YT'SKSSK	SKOSRKS	PGDDEDDKD	170			
126	KGN-AGSSDDEEG-KLV	IDEPAKEKVEKAL	KGRAGD	LLEDSPKPK	EAENPEGEKEAA	183			
171	CKE-----EENKSS	EGGDAGNDTRNT	TSDLQKTS	EGT	203				
184	TLEVRPLMVEVKNKSTP	SGSGRPPQ	DEEBEED	EBEEAT	225				

RESULT 2  
-08-760-745-1  
Sequence 1, Application US/08760745  
Patent No. 5972658  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Goli, Surya K.  
APPLICANT: Murry, Lynn E.  
TITLE OF INVENTION: LUNG GROWTH FACTOR VARIANT  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/760,745  
FILING DATE: Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0169 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 235 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: lungast01  
CLONE: 876242  
08-760-745-1

```

Db      2 PHAFKCDLVFAKMKYGPWHPARIDDIADGAVKPPNKNYIFPFGTHETAFLGPKDLFPY 61
QY      66 KEYKDRFGKSNKRKGFNEGLWEIENNPNPGVKFTGYQAIQQOSSSE-----TEGEG--- 114
Db      62 EESKEKFGPNKRKGFSEGLWEIENNPTVKASGYQSSQKKSCVPEPEPEAAEGDGDKK 121
QY      115 GNTADASSBEEGDRVEDGKKGKNEKAGSKRKKSS---YTSKKSSKOSRKSPGDEDDKDC 171
Db      122 GN-AEGSSDEEG-KLVIDPEAKEKNEKALKRAGDILLEDSPKRPKEAENPEGEKEAAT 179
QY      172 KE-----BENKSSEGG 183
Db      180 LEVERPLPMEVEKNSPTSEPG 200

RESULT 3
US-08-760-745-5
; Sequence 5, Application US/08760745
; Patent No. 5972658
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Murry, Lyvin E.
; TITLE OF INVENTION: LUNG GROWTH FACTOR VARIANT
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,745
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0169 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 945419
US-08-760-745-5

```

```
Query Match      49.0%; Score 532; DB 2; Length 237;  
Best Local Similarity 53.9%; Pred. No. 2.9e-46;  
Matches 111; Conservative 29; Mismatches 38; Indels 28; Gaps 6  
  
Qy    5 RPEYKAGDLVFAMKGYPHPWARIIDELPSGAVKPANKYPIRPFGTHTAFGLPKDLP 64  
       |||||  
Db     6 RQEYKCGDLVFAMKGYPHPWARIDEMPPAAVKSTANKIQVFFGTHTAFGLPKDLP 65  
       |||||  
  
Qy    65 YKEIKDKFGKSNIKRKFNGEGLWIENNPGYKFTGYQAIOQQS-----SSETGEIG 114  
       |||||
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2004, 15:29:12 ; Search time 33 Seconds  
(without alignments)  
1257.869 Million cell updates/sec

Title: US-09-787-328B-4  
Perfect score: 1086  
Sequence: 1 MARPPREYKAGDLVFAKMK.....DAGNDTRTTSIDLQKTSGET 203

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 762491 seqs, 204481190 residues

Total number of hits satisfying chosen parameters: 762491

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1086	100.0	203	9	US-09-788-826-41
2	1086	100.0	203	12	US-10-247-671-132
3	1033.5	97.0	202	11	US-09-866-050A-651
4	895	82.4	172	9	US-09-925-301-987
5	749	69.0	145	9	US-09-768-826-60
6	535.5	49.3	246	9	US-09-925-302-511
7	534.5	49.2	240	9	US-09-938-885A-3
8	534.5	49.2	240	11	US-09-987-755-7
9	534.5	49.2	240	12	US-10-116-275-220
10	534.5	49.2	240	15	US-10-207-791-2
11	532.5	49.0	235	9	US-09-938-885A-1
12	532	49.0	237	9	US-09-938-885A-5
13	524	48.3	314	9	US-09-768-826-43
14	524	48.3	670	11	US-09-823-187-86
15	524	48.3	670	12	US-09-863-776-51

16	524	48.3	671	11	US-09-946-374-308	Sequence 308, App
17	524	48.3	671	11	US-09-823-187-85	Sequence 85, Appl
18	524	48.3	671	12	US-10-015-387A-308	Sequence 308, Appl
19	524	48.3	671	12	US-10-006-130A-308	Sequence 308, App
20	524	48.3	671	12	US-10-199-672-346	Sequence 346, App
21	524	48.3	671	12	US-10-006-172A-308	Sequence 308, App
22	524	48.3	671	12	US-10-187-749-346	Sequence 346, App
23	524	48.3	671	12	US-10-194-457-346	Sequence 346, App
24	524	48.3	671	12	US-10-184-642-346	Sequence 346, App
25	524	48.3	671	12	US-10-196-747-346	Sequence 346, App
26	524	48.3	671	12	US-10-015-392A-308	Sequence 308, App
27	524	48.3	671	12	US-10-017-253A-308	Sequence 308, App
28	524	48.3	671	12	US-10-173-689-346	Sequence 346, App
29	524	48.3	671	12	US-10-173-690-346	Sequence 346, App
30	524	48.3	671	12	US-10-173-691-346	Sequence 346, App
31	524	48.3	671	12	US-10-173-692-346	Sequence 346, App
32	524	48.3	671	12	US-10-173-694-346	Sequence 346, App
33	524	48.3	671	12	US-10-173-698-346	Sequence 346, App
34	524	48.3	671	12	US-10-173-709-346	Sequence 346, App
35	524	48.3	671	12	US-10-173-707-346	Sequence 346, App
36	524	48.3	671	12	US-10-174-569-346	Sequence 346, App
37	524	48.3	671	12	US-10-174-583-346	Sequence 346, App
38	524	48.3	671	12	US-10-174-587-346	Sequence 346, App
39	524	48.3	671	12	US-10-174-589-346	Sequence 346, App
40	524	48.3	671	12	US-10-174-591-346	Sequence 346, App
41	524	48.3	671	12	US-10-175-736-346	Sequence 346, App
42	524	48.3	671	12	US-10-175-742-346	Sequence 346, App
43	524	48.3	671	12	US-10-175-744-346	Sequence 346, App
44	524	48.3	671	12	US-10-175-745-346	Sequence 346, App
45	524	48.3	671	12	US-10-175-748-346	Sequence 346, App

ALIGNMENTS

RESULT 1  
US-09-768-826-41  
; Sequence 41, Application US/09768826  
; Patent No. US20020012966A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi et al.  
; TITLE OF INVENTION: 18 human secreted proteins  
; FILE REFERENCE: PF512P1  
; CURRENT APPLICATION NUMBER: US/09768, 826  
; CURRENT FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: PCT/US00/22350  
; PRIOR FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: 60/148, 759  
; PRIOR FILING DATE: 1999-08-16  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 41  
; LENGTH: 203  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-768-826-41

Query Match	100.0%;	Score 1086;	DB 9;	Length 203;
Best Local Similarity	100.0%;	Pred. No. 4.9e-82;		
Matches 203;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MARPPREYKAGDLVFAKMKGYHPWPARIDELPEGAVKPPANKYPPIFFGTHETAFGLCPK	60	
Db	1	MARPPREYKAGDLVFAKMKGYHPWPARIDELPEGAVKPPANKYPPIFFGTHETAFGLCPK	60	
Qy	61	DLFPYKEYKDKFGSKNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGEGNTADA	120	
Db	61	DLFPYKEYKDKFGSKNKRKGFNEGLWEIENNPVGKFTGYQAIQQSSSETEGEGNTADA	120	
Qy	121	SSEEGDRVEEDGKGRKNEKAGSKRKYTSKKSKQSKRSPGDEDDKCKKEENKSSS	180	
Db	121	SSEEGDRVEEDGKGRKNEKAGSKRKYTSKKSKQSKRSPGDEDDKCKKEENKSSS	180	

```
QY 181 EGGDAGNDTRNTTSDLOKTSEGT 203
Db 181 EGGDAGNDTRNTTSDLOKTSEGT 203

RESULT 2
US-10-247-671-132
; Sequence 132, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 132
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 2676869CD1
US-10-247-671-132

Query Match 100.0%; Score 1086; DB 12; Length 203;
Best Local Similarity 100.0%; Pred. No. 4.9e-82; Indels 0; Gaps 0;
Matches 203; Conservative 0; Mismatches 0;

QY 1 MARPPREYKAGDLVFAKMKGYPHWPARIDELPEGAVKPPANKYPIFFFGTHTAFLGPK 60
Db 1 MARPPREYKAGDLVFAKMKGYPHWPARIDELPEGAVKPPANKYPIFFFGTHTAFLGPK 60

QY 61 DLFPYKEYKDKFGKSNKRGFNEGLWEIENNPVGFYQAIQQSSSSETEGEGNTADA 120
Db 61 DLFPYKEYKDKFGKSNKRGFNEGLWEIENNPVGFYQAIQQSSSSETEGEGNTADA 120

QY 121 SSEEEDRVEEDGKGRKNEKAGSKRKYTSKSKSQSRKSPGDEDDKCKEENKSSS 180
Db 121 SSEEEDRVEEDGKGRKNEKAGSKRKYTSKSKSQSRKSPGDEDDKCKEENKSSS 180

QY 181 EGGDAGNDTRNTTSDLOKTSEGT 203
Db 181 EGGDAGNDTRNTTSDLOKTSEGT 203

RESULT 3
US-09-866-050A-651
; Sequence 651, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 651
; LENGTH: 202
; TYPE: PRT

Query Match 82.4%; Score 895; DB 9; Length 172;
Best Local Similarity 98.8%; Pred. No. 2.3e-66;
Matches 169; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 33 PEGAVKPPANKYPIFFFGTHTAFLGPKDLFPYKEYKDKFGKSNKRGFNEGLWEIENNP 92
Db 2 PRGAVKPSANKYPIFFFGTHTAFLGPKDLFPYKEYKDKFGKSNKRGFNEGLWEIENNP 61

QY 93 GVKFTGYQAIQQSSSSETEGEGNTADASSSEEGDRVVEDGKGRKNEKAGSKRKYTS 152
Db 62 GVKFTGYQAIQQSSSSETEGEGNTADASSSEEGDRVVEDGKGRKNEKAGSKRKYTS 121

QY 153 KSKSQSRKSPGDEDDKCKEENKSSSEGGDAGNDTRNTTSDLOKTSEGT 203
Db 122 KSKSQSRKSPGDEDDKCKEENKSSSEGGDAGNDTRNTTSDLOKTSEGT 172

RESULT 4
US-09-925-301-987
; Sequence 987, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 987
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-301-987

Query Match 82.4%; Score 895; DB 9; Length 172;
Best Local Similarity 98.8%; Pred. No. 2.3e-66;
Matches 169; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 33 PEGAVKPPANKYPIFFFGTHTAFLGPKDLFPYKEYKDKFGKSNKRGFNEGLWEIENNP 92
Db 2 PRGAVKPSANKYPIFFFGTHTAFLGPKDLFPYKEYKDKFGKSNKRGFNEGLWEIENNP 61

QY 93 GVKFTGYQAIQQSSSSETEGEGNTADASSSEEGDRVVEDGKGRKNEKAGSKRKYTS 152
Db 62 GVKFTGYQAIQQSSSSETEGEGNTADASSSEEGDRVVEDGKGRKNEKAGSKRKYTS 121

QY 153 KSKSQSRKSPGDEDDKCKEENKSSSEGGDAGNDTRNTTSDLOKTSEGT 203
Db 122 KSKSQSRKSPGDEDDKCKEENKSSSEGGDAGNDTRNTTSDLOKTSEGT 172

RESULT 5
US-09-768-826-60
; Sequence 60, Application US/09768826
; Patent No. US20020012966A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: 18 human secreted proteins
; FILE REFERENCE: PF512P1
; CURRENT APPLICATION NUMBER: US/09/768,826
; CURRENT FILING DATE: 2001-01-25
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2004, 15:28:22 ; Search time 23 Seconds  
(without alignments)  
648.876 Million cell updates/sec

Title: US-09-787-328B-4

Perfect score: 1086

Sequence: 1 MARPPREYKAGDLVFAKMK.....DAGNDRNTTSDLQKTSECT 203

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 352987 seqs, 73517924 residues

Total number of hits satisfying chosen parameters: 352987

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/prodata/1/paa/PCT\_NEW\_COMB.pcp.\*
- 2: /cgn2\_6/prodata/1/paa/US06\_NEW\_COMB.pcp.\*
- 3: /cgn2\_6/prodata/1/paa/US07\_NEW\_COMB.pcp.\*
- 4: /cgn2\_6/prodata/1/paa/US08\_NEW\_COMB.pcp.\*
- 5: /cgn2\_6/prodata/1/paa/US09\_NEW\_COMB.pcp.\*
- 6: /cgn2\_6/prodata/1/paa/US10\_NEW\_COMB.pcp.\*
- 7: /cgn2\_6/prodata/1/paa/US60\_NEW\_COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	486.5	44.8	530	1	PCT-US03-38193-1115
2	486.5	44.8	530	6	US-10-723-860-1115
3	343	31.6	65	6	US-10-425-114A-41803
4	189	17.4	547	6	US-10-745-586-137
5	189	17.4	553	6	US-10-481-779-12
6	186.5	17.2	475	5	US-09-614-150A-2601
7	134.5	12.4	764	7	US-60-490-890-1847
8	134	12.3	765	6	US-10-679-063-25865
9	129.5	11.9	304	6	US-10-739-930-7468
10	128.5	11.8	764	6	US-10-679-063-7384
11	128.5	11.8	764	6	US-10-679-063-20098
12	127.5	11.7	727	6	US-10-679-063-7103
13	127.5	11.7	727	6	US-10-679-063-15050
14	127.5	11.7	764	6	US-10-679-063-7102
15	127.5	11.7	764	6	US-10-679-063-15049
16	127.5	11.7	1360	1	PCT-US03-36702-24
17	127.5	11.7	1360	6	US-10-714-228-24
18	127.5	11.7	1360	7	US-60-490-890-211
19	124.5	11.5	191	6	US-10-679-063-13279
20	122.5	11.3	488	1	PCT-US03-38193-2384
21	122.5	11.3	488	6	US-10-723-860-2984
22	122.5	11.3	488	7	US-60-485-450-1300
23	122.5	11.3	488	7	US-60-495-135-377
24	122.5	11.3	488	7	US-60-495-135-378
25	122.5	11.3	498	7	US-60-531-203-3
26	122.5	11.3	498	7	US-60-531-567-3

27	122.5	11.3	498	7	US-60-531-371-3	Sequence 3, Appli
28	119.5	11.0	231	5	US-09-614-150A-21075	Sequence 21075, A
29	119.5	11.0	315	6	US-10-425-114A-38959	Sequence 38959, A
30	119.5	11.0	315	6	US-10-425-114A-46993	Sequence 46993, A
31	119.5	11.0	315	6	US-10-425-114A-53526	Sequence 53526, A
32	119.5	11.0	315	6	US-10-425-114A-60313	Sequence 60313, A
33	119	11.0	1278	5	US-09-614-150A-42438	Sequence 42438, A
34	118	10.9	312	6	US-10-425-114A-47459	Sequence 47459, A
35	118	10.9	517	6	US-10-425-114A-62674	Sequence 62674, A
36	118	10.9	554	6	US-10-425-114A-42986	Sequence 42986, A
37	117.5	10.8	637	6	US-10-679-063-18043	Sequence 18043, A
38	116.5	10.7	422	6	US-10-739-930-9013	Sequence 9013, Ap
39	116	10.7	281	6	US-10-296-115-918	Sequence 918, App
40	116	10.7	1664	5	US-09-614-150A-2484	Sequence 2484, Ap
41	114	10.5	629	7	US-60-490-890-1888	Sequence 1888, Ap
42	114	10.5	802	7	US-60-490-890-1890	Sequence 1890, Ap
43	114	10.5	1365	7	US-60-490-890-1892	Sequence 1892, Ap
44	113	10.4	658	7	US-60-490-890-2143	Sequence 2143, Ap
45	112	10.3	254	6	US-10-425-114A-61267	Sequence 61267, A

ALIGNMENTS

RESULT 1

PCT-US03-38193-1115

; Sequence 1115, Application PC/TUS0338193

; GENERAL INFORMATION:

; APPLICANT: Ariz, Natasha

; APPLICANT: Ginsburg, Wendy M.

; APPLICANT: Zlotnik, Albert

; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &

; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators

; FILE REFERENCE: 05882.0193.00PC00

; CURRENT APPLICATION NUMBER: PCT/US03/38193

; CURRENT FILING DATE: 2003-11-26

; PRIOR APPLICATION NUMBER: 60/429,739

; PRIOR FILING DATE: 2002-11-26

; NUMBER OF SEQ ID NOS: 8393

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1115

; LENGTH: 530

; TYPE: PRT

; ORGANISM: Homo sapiens

PCT-US03-38193-1115

Query Match 44.8%; Score 486.5; DB 1; Length 530;

Best Local Similarity 38.8%; Pred. No. 3.6e-24;

Matches 109; Conservative 29; Mismatches 42; Indels 101; Gaps 7;

QY 7 REYKAGDLVFAKMGYPHWPARIDELPEGAVKPPANKYPIFPFGTHETAPLGPDLFPYK 66

Db 3 RUFKPGDLIFAKMGYPHWPARVDEPDGAVKPPANKLPFPFGTHETAPLGPDLFPYK 62

QY 67 EYKDFGKSNKRKGFNEGLWEIENPNPGVKTGYQAIQQQS--SSETEGEGGNTA----- 118

Db 63 ENKEYGKPKRKGNEGLWEIDNNPKVFKSSQQAATKQSNASSDVEVEKETSVSKEDT 122

QY 119 -----DASSEE-----EGDRVEEDGKGRK--NEKAGSKR--KKSVTSKSSK 157

Db 123 DHEEKASNEEDTVKAVDITTPKAARRKRKAQVETEEAGVVVTTATASVNLKVSPPGR 182

QY 125 -----EGDRVEEDGKGRK--NEKAGSKR--KKSVTSKSSK 157

Db 183 PAATEVKIPKPRGRPMVKQPCPSSSDIITEEDKSKKQGEERKQPKPKQKEEGKEED 242

QY 158 QSRKSPGDEDDKCKEENK-----SSSEGGDAGND 188

Db 243 KPRKEPDKKEGK--KEVESKRNKLAKTGVTSTSDSEEGDD 281

RESULT 2

US-10-723-860-1115



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2004, 15:26:52 ; Search time 21 Seconds

(without alignments)

929.631 Million cell updates/sec

Title: US-09-787-328B-4

Perfect score: 1086

Sequence: 1 MARPRREYKAGDLVFAKMK.....DAGNDTRNTTSDLQKTSEGT 203

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1086	100.0	203	JC7163	hepatoma-derived g
2	534.5	49.2	240	A55055	hepatoma-derived g
3	532	49.0	237	JC5660	hepatoma-derived g
4	504.5	46.5	669	JC5662	hepatoma-derived g
5	486.5	44.8	530	JC7168	lens epithelium-de
6	337.5	31.1	283	JC5661	hepatoma-derived g
7	193	17.8	1095	JC25520	hypothetical prote
8	162.5	15.0	1392	T51947	probable transcrip
9	162	14.9	1445	T50508	hypothetical prote
10	161	14.8	1366	B84924	hypothetical prote
11	155.5	14.3	1261	T48096	hypothetical prote
12	142	13.1	568	T39897	hypothetical prote
13	134.5	12.4	420	T01077	hypothetical prote
14	134.5	12.4	727	S18193	autoantigen NOR-90
15	134.5	12.4	764	S09318	transcription fact
16	134	12.3	765	S22314	transcription fact
17	132	12.2	488	I46014	cylindrin II - bovin
18	131	12.1	524	S35551	transcription fact
19	130	12.0	210	T28771	hypothetical prote
20	130	12.0	211	T25911	hypothetical prote
21	128.5	11.8	727	B40439	UBF transcription
22	128.5	11.8	764	A40439	UBF transcription
23	128.5	11.8	1093	T50652	AP-3 complex beta3
24	128.5	11.8	1094	T50651	AP-3 complex beta3
25	127.5	11.7	727	JC5113	ribosomal transcri
26	127.5	11.7	764	JC5112	ribosomal transcri
27	127.5	11.7	1360	JC5839	GRBP-N protein - h
28	127	11.7	532	T06029	hypothetical prote
29	124.5	11.5	1305	T00670	probable inositol

30	122.5	11.3	522	2	C96608	hypothetical prote
31	121.5	11.2	307	2	T04141	histone deacetylase
32	121.5	11.2	902	2	T47966	hypothetical prote
33	120.5	11.1	219	2	T19897	hypothetical prote
34	120.5	11.1	598	2	B40713	cylindrin I - human
35	119.5	11.0	2206	2	G71611	hypothetical prote
36	118.5	10.9	751	2	S65469	DNA topoisomerase
37	118.5	10.9	806	2	T23648	hypothetical prote
38	118.5	10.9	1079	2	T38913	translation initia
39	117.5	10.8	327	2	T30072	hypothetical prote
40	117.5	10.8	414	2	A39205	nuclear localizati
41	117.5	10.8	1428	1	ISBYT2	DNA topoisomerase
42	117	10.8	654	2	B71623	knob-associated Hi
43	117	10.8	763	2	T08929	hypothetical prote
44	116	10.7	697	2	T16908	hypothetical prote
45	115.5	10.6	1178	2	S78475	mannosylphosphoryl

ALIGNMENTS

RESULT 1

JC7163

hepatoma-derived growth factor-related protein 3 - human

C;Species: Homo sapiens (man)

C;Date: 04-Mar-2000 #sequence\_revision 04-Mar-2000 #text\_change 21-Jul-2000

C;Accession: JC7163

R;Kegame, K.; Yamamoto, M.; Kishima, Y.; Enomoto, H.; Yoshida, K.; Suemura, M.; Kishim

Biochem. Biophys. Res. Commun. 266, 81-87, 1999

A;Title: A new member of a hepatoma-derived growth factor gene family can translocate t

A;Reference number: JC7163; MUID:20050055; PMID:10581169

A;Accession: JC7163

A;Molecule type: mRNA

A;Residues: 1-203 <IKE>

A;Cross-references: DDBJ:AB029156; NID:96855467; PIDN:BAA90477.1; PID:96855468

C;Comment: This protein, a new member of the hepatoma-derived growth factor gene family

associates to the nucleus and induces cellular growth.

C;Genetics:

A;Gene: hrp-3

A;Map position: 15q25

C;Keywords: differentiation; glycoprotein; growth factor; nucleus

Query Match 100.0%; Score 1086; DB 2; Length 203;

Best Local Similarity 100.0%; Pred. No. 2,7e-69;

Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARPRREYKAGDLVFAKMKGYPHWPARIDELPEGAVKPPANKYPIFFGTHETAFILGPK 60

Db 1 MARPRREYKAGDLVFAKMKGYPHWPARIDELPEGAVKPPANKYPIFFGTHETAFILGPK 60

Qy 61 DLPPYKEYDKPKSKNRKGFNEGLWEIENNPVGKVTGYQAIQQSSSETEGEGGNTADA 120

Db 61 DLPPYKEYDKPKSKNRKGFNEGLWEIENNPVGKVTGYQAIQQSSSETEGEGGNTADA 120

Qy 121 SSEEEDRVEEDGKGRKNEKAGSKRKKSVTSKKSKQSRKSPGDDDDKCKEENKSSS 180

Db 121 SSEEEDRVEEDGKGRKNEKAGSKRKKSVTSKKSKQSRKSPGDDDDKCKEENKSSS 180

Qy 181 EGGDAGNDTRNTTSDLQKTSEGT 203

Db 181 EGGDAGNDTRNTTSDLQKTSEGT 203

RESULT 2

A55055

hepatoma-derived growth factor - human

C;Species: Homo sapiens (man)

C;Date: 18-Nov-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Jun-2000

C;Accession: A55055

R;Nakamura, H.; Izumoto, Y.; Kambe, H.; Kuroda, T.; Mori, T.; Kawamura, K.; Yamamoto, H

J. Biol. Chem. 269, 25143-25149, 1994

A;Title: Molecular cloning of complementary DNA for a novel human hepatoma-derived grow

A;Reference number: A55055; MUID:95014294; PMID:7929202

```
A:Accession: A55055
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-240 <NAK>
A:Cross-references: GB:D16431; NID:G598955; PIDN:BAA03903.1; PID:G598956
C:Genetics:
A:Gene: GDB:HDGF
A:Cross-references: GDB:437554
A:Map position: Xq25-Xq25
C:Keywords: heparin binding

Query Match      49.2%; Score 534.5; DB 2; Length 240;
Best Local Similarity 41.4%; Pred. No. 1.3e-30;
Matches 114; Conservative 25; Mismatches 58; Indels 25; Gaps 6;

QY 5 RREYKAGDLVFAKMGYPHWPARIDELPEGAVKPPANKYPIFFFGTHTAFLGPKDLFP 64
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6 ROKEYKCGDLVFAKMGYPHWPARIDEMPEAAVKSTANKYQVFFFGTHTAFLGPKDLFP 65
QY 65 YKEYKDKFKGSKNRKGFNEGLWEIENNPVGKFTGYQAIQQQSSSE-----TEGEG-- 114
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 YYESKEKFKPKNRKGFSGLEWEIENNPVTKASGYQSSQKSCVPEPEPEAAEGDGDK 125
QY 115 -GNTADASSEEGDRVEEDGKGRKNEKAGSKRKKS---YTSKSKSKSRKSPGDEDDKD 170
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
126 KGN-AEGSSDEEG-KLVIDEPAKNEKAGALKRRAGDLDSPKRPKEAENPEGEKEAA 183
QY 171 CKE-----EENKSSSEGGDAGNDTRNTTSDLOKTSGET 203
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
184 TLEVERPLMEVEKKNSTPSEPGSGRGPQOESEEEDEBEAT 225

RESULT 3
JC5660
hepatoma-derived growth factor - mouse
N:Alternate names: HDGF
C:Species: Mus musculus (house mouse)
C:Date: 17-Nov-1997 #sequence_revision 17-Nov-1997 #text_change 17-Mar-1999
C:Accession: JC5660
R:izumoto, Y.; Kuroda, T.; Harada, H.; Kishimoto, T.; Nakamura, H.
Biochem. Biophys. Res. Commun. 238, 26-32, 1997
A:Title: Hepatoma-derived growth factor belongs to a gene family in mice showing significant
A:Reference number: JC5660; MUID:97445118; PMID:9299445
A:Molecule type: mRNA
A:Residues: 1-237 <IZU>
A:Cross-references: DDBJ:D63707
C:Comment: This protein translocates to the nucleus and directly functions in mitogenic
F:1-98/Region: hath #status predicted
F:155-170/Region: nuclear location signal

Query Match      49.0%; Score 532; DB 2; Length 237;
Best Local Similarity 53.9%; Pred. No. 1.9e-30;
Matches 111; Conservative 29; Mismatches 38; Indels 28; Gaps 6;

QY 5 RREYKAGDLVFAKMGYPHWPARIDELPEGAVKPPANKYPIFFFGTHTAFLGPKDLFP 64
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6 ROKEYKCGDLVFAKMGYPHWPARIDEMPEAAVKSTANKYQVFFFGTHTAFLGPKDLFP 65
QY 65 YKEYKDKFKGSKNRKGFNEGLWEIENNPVGKFTGYQAIQQQSS-----SSETEGEG 114
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 YYESKEKFKPKNRKGFSGLEWEIENNPVTKASGYQSSQKSCAAEPEPEPEAAEGDGDK 125
QY 115 GNTADASSEEGDRVEEDGKGRKNEKAGSKRKKSYSKSKSKSRKSPGDEDDKDCKE- 173
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
126 KGSAGSSDEEG-KLVIDEPAKNEKAGTLEKRRAGDVLEDSKRPKES-GDHEED-KEI 182
QY 174 -----EENKSSSEGGDAG 186
Db 183 AALEGERPLPVEVEKNSTPSE-PSDG 207

RESULT 4
JC5662
hepatoma-derived growth factor-related protein 2 - mouse
N:Alternate names: HRP-2
C:Species: Mus musculus (house mouse)
C:Date: 17-Nov-1997 #sequence_revision 17-Nov-1997 #text_change 05-Nov-1999
C:Accession: JC5662
R:izumoto, Y.; Kuroda, T.; Harada, H.; Kishimoto, T.; Nakamura, H.
Biochem. Biophys. Res. Commun. 238, 26-32, 1997
A:Title: Hepatoma-derived growth factor belongs to a gene family in mice showing significant
A:Reference number: JC5660; MUID:97445118; PMID:9299445
A:Accession: JC5662
A:Molecule type: mRNA
A:Residues: 1-669 <IZU>
A:Cross-references: DDBJ:D63850; NID:G2558500; PIDN:BAA22896.1; PID:d1023766; PID:G25585
C:Comment: This protein translocates to the nucleus and directly functions in mitogenic
F:1-93/Region: hath #status predicted
F:321-363/Region: nuclear location signal

Query Match      46.5%; Score 504.5; DB 2; Length 669;
Best Local Similarity 46.9%; Pred. No. 4.8e-28;
Matches 100; Conservative 26; Mismatches 62; Indels 25; Gaps 3;

QY 6 PREYKAGDLVFAKMGYPHWPARIDELPEGAVKPPANKYPIFFFGTHTAFLGPKDLFP 65
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2 PHAFKPGDLVFAKMGYPHWPARIDDIADGAVKPPNPKYPIFFFGTHTAFLGPKDLFP 61
QY 66 KEYKDKFKGSKNRKGFNEGLWEIENNPVGKFTGYQAIQQQSSSETEGEGNTADASSEEE 125
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
62 DKCKDKYKPKNRKGFNEGLWEIQNNPHASYSAPPPVSSSSDSEAPEADLGCSDVDKDK 121
QY 126 -----GDRVEEDGKGRKNEKAGSKRKKSYSKSKSKSRKSPGDEDDKD 170
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
122 SRRVMTVTAVTTATSDRMESDSDSKSDSGSLKRRKTPVLKVSVKARRASSDLDOAS 181
QY 171 C--KEENKSSSEGGDAGNDTRNTTSDLOKTSSE 201
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182 VSPSEDSSESPSES-----EXTSDQDFTPE 206

RESULT 5
JC7168
lens epithelium-derived growth factor - human
N:Alternate names: transcriptional co-activator p75 protein
C:Species: Homo sapiens (man)
C:Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 11-May-2000
C:Accession: JC7168
R:Singh, D.P.; Ohguro, N.; Kikuchi, T.; Sueno, T.; Reddy, V.N.; Yuge, K.; Chylack Jr., I.
Biochem. Biophys. Res. Commun. 267, 373-381, 2000
A:Title: Lens epithelium-derived growth factor: Effects on growth and survival of lens e
A:Reference number: JC7168; MUID:20090645; PMID:10623627
A:Accession: JC7168
A:Molecule type: mRNA
A:Residues: 1-530 <SIN>
A:Cross-references: GB:AF063020
A:Experimental source: lens epithelial cells
C:Comment: This protein, a regulatory factor, belongs to a family of growth factor, whic
on of this factor stimulates cell attachment, shortenes the initial lag in growth, and e
C:Genetics:
A:Gene: ledgf
C:Keywords: eye lens; glycoprotein; growth factor; growth regulation

Query Match      44.8%; Score 486.5; DB 2; Length 530;
Best Local Similarity 38.8%; Pred. No. 6.8e-27;
Matches 109; Conservative 29; Mismatches 42; Indels 101; Gaps 7;

QY 7 REYKAGDLVFAKMGYPHWPARIDELPEGAVKPPANKYPIFFFGTHTAFLGPKDLFP 66
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3 RDKFPGDLVFAKMGYPHWPARVDEPDGAVKPPNKLPIFFFGTHTAFLGPKDLFP 62
QY 67 BYKDKFKGSKNRKGFNEGLWEIENNPVGKFTGYQAIQQQSS-----SSETEGEGNTA----- 118
Db 63 ENKEKYKPKNRKGFNEGLWEIDNNPKVKFSSQQAATQSNASSDVEVEEKTSVKEDT 122
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2004, 15:26:51 ; Search time 18 seconds  
(without alignments)

530.357 Million cell updates/sec

Title: US-09-787-328B-4

Perfect score: 1086

Sequence: 1 MARPREYKAGDLVFAKMK.....DAGNDTRNTSDLQKTSEGT 203

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	534.5	49.2	240	1 HDGF_HUMAN	P51858 homo sapien
2	532	49.0	237	1 HDGF_MOUSE	P51859 mus musculus
3	134.5	12.4	764	1 UBFI_HUMAN	P17480 homo sapien
4	134	12.3	765	1 UBFI_MOUSE	P25976 mus musculus
5	132	12.2	488	1 T2FA_BOVIN	Q28092 bos taurus
6	131	12.1	524	1 T2FA_XENLA	Q04870 xenopus lae
7	128.5	11.8	764	1 UBFI_RAT	P25977 rattus norv
8	128.5	11.8	1094	1 A3BI_HUMAN	Q00203 homo sapien
9	127.5	11.7	406	1 NSBI_MOUSE	P52701 homo sapien
10	127.5	11.7	1360	1 MSH6_HUMAN	Q99mf6 mus musculus
11	126.5	11.6	875	1 ARS2_MOUSE	P54276 mus musculus
12	124	11.4	1358	1 MSH6_MOUSE	Q92769 homo sapien
13	122.5	11.3	488	1 HD42_HUMAN	P35663 homo sapien
14	120.5	11.1	598	1 CYLI_HUMAN	Q10251 saccharomyc
15	118.5	10.9	1079	1 IFP2_SCHPO	P27476 saccharomyc
16	117.5	10.8	414	1 NSRI_YEAST	Q28181 bos taurus
17	117.5	10.8	1394	1 CNG4_BOVIN	P06786 saccharomyc
18	117.5	10.8	1428	1 TOP2_YEAST	P56519 gallus gall
19	116.5	10.7	488	1 HD42_CHICK	Q96b71 homo sapien
20	116	10.7	556	1 HIR3_HUMAN	P41848 caenorhabdi
21	116	10.7	697	1 SSRP_CABEL	P36044 saccharomyc
22	115.5	10.6	1178	1 MN4_YEAST	Q92pp1 berberis st
23	113.5	10.5	416	1 CRIC_BERST	Q13367 homo sapien
24	113.5	10.5	1082	1 A3B2_HUMAN	P41891 schizosacch
25	113	10.4	500	1 GAR2_SCHPO	Q01033 herpesvirus
26	113	10.4	676	1 S521_HUMAN	P13666 oryctolagus
27	112.5	10.4	797	1 VG48_HSVSA	Q42227 xenopus lae
28	112.5	10.4	908	1 SRCA_RABIT	Q13547 homo sapien
29	111.5	10.3	480	1 HD12_XENLA	P70288 mus musculus
30	111.5	10.3	482	1 HD42_HUMAN	P40631 tetrahymena
31	111.5	10.3	488	1 HD42_MOUSE	Q921t1 mus musculus
32	111.5	10.3	633	1 MLH_TETTH	
33	111.5	10.3	1105	1 A3BI_MOUSE	

## RESULT 1

ID	HDGF_HUMAN	STANDARD;	PRT;	240 AA.
AC	P51858;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Hepatoma-derived growth factor (HDGF) (High-mobility group protein 1-like 2) (HMG-1L2).			
GN	HDGF OR HMG1L2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 4-24.			
RC	TISSUE=Hepatoma;			
RX	MEDLINE=95014294; PubMed=7929202;			
RA	Nakamura H., Izumoto Y., Kambe H., Kuroda T., Mori T., Kawamura K., Yamamoto H., Kishimoto T.;			
RT	"Molecular cloning of complementary DNA for a novel human hepatoma-derived growth factor. Its homology with high mobility group-1 protein."			
RT	J. Biol. Chem. 269:25143-25149 (1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E., Schnarch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).			
CC	-!- FUNCTION: Heparin-binding protein, with mitogenic activity for fibroblasts.			
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic.			
CC	-!- TISSUE SPECIFICITY: Ubiquitous.			
CC	-!- SIMILARITY: Contains 1 FWWP domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -			

SEQUENCE FROM N. A.  
STRAIN=C57BL/6J; TISSUE=Embryo;  
MEDLINE=21035660; PubMed=11217851;  
Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I., Saito T., Okazaki Y., Gotjohori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavante T., Frieschmann W., Gaasterland T., Gissi C., King B., Kochiya H., Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J., Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald W., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;  
"Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001).

[4] SEQUENCE FROM N.A.  
TISSUE=Liver, and Mammary gland;  
MEDLINE=22398257; PubMed=12477932;  
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,  
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,  
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
"Genome and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences.",  
-!- FUNCTION: Heparin-binding protein, with mitogenic activity for fibroblasts (By similarity).  
-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
-!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN TESTIS AND SKELETAL MUSCLE, TO INTERMEDIATE EXTENTS IN HEART, BRAIN, LUNG, LIVER, AND KIDNEY, AND TO A MINIMAL EXTENT IN SPLEEN.  
-!- SIMILARITY: Contains 1 PWWP domain.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL: D63707; BAA09838.1; - -  
EMBL: AF251787; AAF65469.1; - -  
EMBL: AK017863; BAB30979.1; - -  
EMBL: BC005713; AAH05713.1; - -  
EMBL: BC021654; AAH21654.1; - -  
PIR: JC5660; JCS5660.  
MGD: MGI:1194494; Hdgf.  
InterPro: IPR000313; PWWP\_domain.  
Pfam: PF00855; PWWP; 1.  
SMART: SM00293; PWWP; 1.

[illegible]



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2004, 15:26:51 ; Search time 35 Seconds  
(without alignments)

1496.705 Million cell updates/sec

Title: US-09-787-328B-4

Perfect score: 1086

Sequence: 1 MARPREYKAGDLVPAKMK.....DAGNDTRNTSDLQKTSEGT 203

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL 23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1086	100.0	203	4 Q9Y3E1	Q9Y3E1 homo sapien
2	1059.5	97.6	202	11 Q9JMG7	Q9JMG7 mus musculus
3	1053.5	97.0	202	11 Q923W4	Q923W4 rattus norv
4	1050.5	96.7	202	11 Q9D2M7	Q9D2M7 mus musculus
5	1045.5	96.3	205	11 Q8BQ69	Q8BQ69 mus musculus
6	851.5	78.4	220	11 Q8BR62	Q8BR62 mus musculus
7	532	49.0	237	11 Q8BPG7	Q8BPG7 mus musculus
8	530.5	48.8	239	6 Q9XSK7	Q9XSK7 bos taurus
9	530	48.8	237	11 Q8VHK7	Q8VHK7 rattus norv
10	530	48.8	237	11 Q923W3	Q923W3 rattus norv
11	524	48.3	670	4 Q9BW08	Q9BW08 homo sapien
12	512.5	47.2	669	11 Q925G1	Q925G1 rattus norv
13	504.5	46.5	669	11 Q35540	Q35540 mus musculus
14	504.5	46.5	678	11 Q99L92	Q99L92 mus musculus
15	486.5	44.8	333	4 Q9N2I3	Q9N2I3 homo sapien
16	486.5	44.8	530	4 Q75475	Q75475 homo sapien

17	485.5	44.7	333	4	Q95368	Q95368 homo sapien
18	485.5	44.7	530	4	Q9UER6	Q9UER6 homo sapien
19	485	44.7	331	11	Q99JF7	Q99JF7 mus musculus
20	485	44.7	528	11	Q95JF8	Q95JF8 mus musculus
21	481	44.3	283	13	Q8AVR7	Q8AVR7 xenopus lae
22	467.5	43.0	146	11	Q9CT03	Q9CT03 mus musculus
23	411	37.8	235	6	Q9XSK6	Q9XSK6 bos taurus
24	361.5	33.3	251	4	Q9NUE8	Q9NUE8 homo sapien
25	354.5	32.6	251	4	Q96MJ6	Q96MJ6 homo sapien
26	342.5	31.5	302	4	Q8NPR8	Q8NPR8 homo sapien
27	342	31.5	296	11	Q8VIH2	Q8VIH2 rattus norv
28	341	31.4	283	11	Q35539	Q35539 mus musculus
29	269.5	24.8	483	6	Q8MJG1	Q8MJG1 bos taurus
30	251.5	23.2	258	11	Q9ESF6	Q9ESF6 mus musculus
31	238	21.9	50	4	Q8N4N4	Q8N4N4 homo sapien
32	193	17.8	832	5	Q95Q90	Q95Q90 caenorhabdi
33	193	17.8	835	5	Q95Q91	Q95Q91 caenorhabdi
34	189	17.4	547	4	Q9BKK2	Q9BKK2 homo sapien
35	188.5	17.4	546	11	Q922P9	Q922P9 mus musculus
36	186.5	17.2	475	5	Q9VAA9	Q9VAA9 drosophila
37	179.5	16.5	546	11	Q9CYQ1	Q9CYQ1 mus musculus
38	167	15.4	523	4	Q9BT11	Q9BT11 homo sapien
39	162.5	15.0	1392	10	Q9XER9	Q9XER9 arabidopsis
40	162	14.9	1445	10	Q9LEY4	Q9LEY4 arabidopsis
41	161	14.8	1366	10	Q8GYT7	Q8GYT7 arabidopsis
42	160.5	14.8	1449	10	Q8H3E3	Q8H3E3 oryza sativ
43	160	14.7	33	11	Q8C9E2	Q8C9E2 mus musculus
44	155.5	14.3	1261	10	Q9LYB5	Q9LYB5 arabidopsis
45	142	13.1	568	3	Q94312	Q94312 achizosacch

#### ALIGNMENTS

#### RESULT 1

Q9Y3E1 ID Q9Y3E1 PRELIMINARY; PRT; 203 AA.  
AC Q9Y3E1;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE CGI-142 protein (Hypothetical protein FLJ10418) (HRP-3) (Hepatoma-derived growth factor 2).  
GN HRP-3 OR HDGF2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP LIN W.-C.;  
RT "Comparative gene cloning: Identification of novel human genes with Caenorhabditis elegans proteome as template."  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Isogai T., Ota T., Hayaashi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., Sugawara M., Takahashi M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Takiguchi S., Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Ninomiya K., Iwayanagi T.;  
RT "NEDO human cDNA sequencing project."  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=20050055; PubMed=10581169;  
RX Ikegami K., Yamamoto M., Kishima Y., Enomoto H., Yoshida K., Suenura M., Kishimoto T., Nakamura H.;  
RT "A new member of a hepatoma-derived growth factor gene family can translocate to the nucleus."  
RL Biochem. Biophys. Res. Commun. 266:81-87(1999).

```
[4]
RN SEQUENCE FROM N.A.
RC TISSUE=uterus;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RA Yu L., Fu Q., Tu Q.;
RT Cloning of a novel human cDNA which is a homolog to mouse hepatoma-
RT derived growth factor (mHGF) and termed hHGF2.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF151900; RAD34137.1; -
DR EMBL; AK001280; BAA91597.1; -
DR EMBL; AB029156; BAA90477.1; -
DR EMBL; BC015483; AAH15483.1; -
DR EMBL; AF110642; AA027003.1; -
DR InterPro; IPR000313; PWMP_domain.
DR Pfam; PF00855; PWMP; 1.
DR PROSITE; PS50812; PWMP; 1.
KW Hypothetical protein.
SQ SEQUENCE 203 AA; 22620 MW; 0B660D665F01659C CRC64;

Query Match 100.0%; Score 1086; DB 4; Length 203;
Best Local Similarity 100.0%; Pred. No. 6.7e-76;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARPPREYKAGDLVFAKMGYPHWPARIDELPEGAVKPPANKYPPIFFGTHETAFLGPK 60
DB 1 MARPPREYKAGDLVFAKMGYPHWPARIDELPEGAVKPPANKYPPIFFGTHETAFLGPK 60

QY 61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGFYQAIQQSSSETEGEGNTADA 120
DB 61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGFYQAIQQSSSETEGEGNTADA 120

QY 121 SSEEGRDVEEDGKGRKNEKAGSKRKYTSKSSKOSRSPGDEDDKCKEENKSSS 180
DB 121 SSEEGRDVEEDGKGRKNEKAGSKRKYTSKSSKOSRSPGDEDDKCKEENKSSS 180

QY 181 EGGDAGNDRNTTSDLOKTSEGT 203
DB 181 EGGDAGNDRNTTSDLOKTSEGT 203

RESULT 2
Q9JMG7 PRELIMINARY; PRT; 202 AA.
AC Q9JMG7;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Hepatoma-derived growth factor-related protein HRP-3.
GN HDGFRP3 OR HRP-3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Ikegama K., Nakamura H.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=20050055; PubMed=10581169;
RA Ikegama K., Yamamoto M., Kishina Y., Enomoto H., Yoshida K.,
RA Sueura M., Kishimoto T., Nakamura H.;
RT "A new member of a hepatoma-derived growth factor gene family can
RT translocate to the nucleus.";
RL Biochem. Biophys. Res. Commun. 266:81-87 (1999).
DR EMBL; AB029493; BAA90478.1; -
DR MGI; MGI:1352760; Hdgfrp3.
DR InterPro; IPR000313; PWMP_domain.
```

```
DR Pfam; PF00855; PWMP; 1.
DR SMART; SM00293; PWMP; 1.
DR PROSITE; PS50812; PWMP; 1.
SQ SEQUENCE 202 AA; 22490 MW; FCDB31A0F79AB8D6 CRC64;

Query Match 97.6%; Score 1059.5; DB 11; Length 202;
Best Local Similarity 98.0%; Pred. No. 7.2e-74;
Matches 199; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 MARPPREYKAGDLVFAKMGYPHWPARIDELPEGAVKPPANKYPPIFFGTHETAFLGPK 60
DB 1 MARPPREYKAGDLVFAKMGYPHWPARIDELPEGAVKPPANKYPPIFFGTHETAFLGPK 60

QY 61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGFYQAIQQSSSETEGEGNTADA 120
DB 61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGFYQAIQQSSSETEGEGNTADA 120

QY 121 SSEEGRDVEEDGKGRKNEKAGSKRKYTSKSSKOSRSPGDEDDKCKEENKSSS 180
DB 121 SSEEGRDVEEDGKGRKNEKAGSKRKYTSKSSKOSRSPGDEDDKCKEENKSSS 180

QY 181 EGGDAGNDRNTTSDLOKTSEGT 203
DB 181 EGGDAGNDRNTTSDLOKTSEGT 203

RESULT 3
Q923W4 PRELIMINARY; PRT; 202 AA.
AC Q923W4;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE HRP3.
GN Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]_TaxID=10116;
RP SEQUENCE FROM N.A.
RA Yu L.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF389347; AAK72965.1; -
DR InterPro; IPR000313; PWMP_domain.
DR Pfam; PF00855; PWMP; 1.
DR SMART; SM00293; PWMP; 1.
DR PROSITE; PS50812; PWMP; 1.
SQ SEQUENCE 202 AA; 22446 MW; FCCDEBBD2BEAB8D6 CRC64;

Query Match 97.0%; Score 1053.5; DB 11; Length 202;
Best Local Similarity 97.5%; Pred. No. 2.1e-73;
Matches 198; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 MARPPREYKAGDLVFAKMGYPHWPARIDELPEGAVKPPANKYPPIFFGTHETAFLGPK 60
DB 1 MARPPREYKAGDLVFAKMGYPHWPARIDELPEGAVKPPANKYPPIFFGTHETAFLGPK 60

QY 61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGFYQAIQQSSSETEGEGNTADA 120
DB 61 DLFPYKEYKDKFGKSNKRKGFNEGLWEIENNPVGFYQAIQQSSSETEGEGNTADA 120

QY 121 SSEEGRDVEEDGKGRKNEKAGSKRKYTSKSSKOSRSPGDEDDKCKEENKSSS 180
DB 121 SSEEGRDVEEDGKGRKNEKAGSKRKYTSKSSKOSRSPGDEDDKCKEENKSSS 180

QY 181 EGGDAGNDRNTTSDLOKTSEGT 203
DB 181 EGGDAGNDRNTTSDLOKTSEGT 203

RESULT 4
Q9D2M7
```



Qy	1	MARPRPREYKAGDLVFAKMGYPHWP	PARIDELPEGAVKPPANKYPIFFGTHETAFLGPK	60
Db	1	MARPRPREYKAGDLVFAKMGYPHWP	PARIDELPEGAVKPPANKYPIFFGTHETAFLGPK	60
Qy	61	DLFPYKEYKDFGKGNKRGKGFNEGLWEI	NNPVGKFTGYQAIQQSSSETEGEGGNTADA	120
Db	61	DLFPYKEYKDFGKGNKRGKGFNEGLWEI	NNPVGKFTGYQTIQQSSSETEGEGGNTADA	120
Qy	121	SSEEGDRVVEEDGKGRKNEKAGSKRK	SKSYTSKKSKSRKSPGDEDDKCKEENKSS	179
Db	121	SSEEGDRV-EDGKGRKNEKGGSKRK	SKSYTSKKSKSRKSOETKMIKTAKKRRTKAA	178
RESULT 7				
Q8BPG7		PRELIMINARY;	PRT; 237 AA.	
AC	Q8BPG7;			
DT	01-MAR-2003 (TrEMBLrel. 23, Created)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)			
DE	Hepatoma-derived growth factor.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Body;			
RX	MEDLINE=22354683; PubMed=12466851;			
RA	The FANTOM Consortium,			
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;			
RT	"Analysis of the mouse transcriptome based on functional annotation of			
RT	60,770 full-length cDNAs";			
RL	Nature 420:563-573 (2002).			
DR	EMBL; AK076021; BAC36126.1; --			
SQ	SEQUENCE 237 AA; 26270 MW; BBA0CF574DA733F CRC64;			
Query Match	49.0%;	Score 532;	DB 11; Length 237;	
Best Local Similarity	53.9%;	Pred. No. 2.5e-33;		
Matches 111;	Conservative 29;	Mismatches 38;	Indels 28; Gaps 6;	
Qy	5	RPRYKAGDLVFAKMGYPHWP	PARIDELPEGAVKPPANKYPIFFGTHETAFLGPKDLFP	64
Db	6	RQKEYKCGDLVFAKMGYPHWP	PARIDEMPEAAVKSTANKYQVFFGTHETAFLGPKDLFP	65
Qy	65	YKEYKDFGKGNKRGKGFNEGLWEI	NNPVGKFTGYQAIQQSS-----SSETEGEG	114
Db	66	YEEKKFKGNKRGKGFSEGLWEI	NNPVTAVGASQSSQKSCAAAEPEVEPEAHEGDK	125
Qy	115	GNTADASSSEEGDRVVEEDGKGRKNEKAGSKRK	SKSYTSKKSKSRKSPGDEDDKCKE- 173	
Db	126	KGSAEGSDSEG-KLVIDEPKERNKEKTLKRRAGDVL	EDSPKPKES-GDHEED-KEI 182	
Qy	174	-----EENKSSSEGGDAG 186		
Db	183	AALGERPLPVEVKNSTPSE-PDSG 207		
RESULT 8				
Q9XSK7		PRELIMINARY;	PRT; 239 AA.	
ID	Q9XSK7			
AC	Q9XSK7;			
DT	01-NOV-1999 (TrEMBLrel. 12, Created)			
DT	01-NOV-1999 (TrEMBLrel. 12, Last sequence update)			
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)			
DE	Hepatoma derived growth factor.			
GN	HDGF.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			